



Health Evidence Review Commission's Value-based Benefits Subcommittee

Part 2

November 13, 2014

8:30 AM

**Meridian Park Hospital
Community Health Education Center, Room 117B&C
19300 SW 65th Avenue, Tualatin, OR 97062**

Section 4.0

New Discussion Items

Coverage of Tonsillar Hypertrophy

Question: What is the HERC intent for coverage for tonsillar hypertrophy

Question source: Lori Andrews, Referral Coordinator at Treasure Valley Pediatric Clinic in Ontario, Oregon

Issue: Tonsillar hypertrophy [ICD-9 code 474.11 (Hypertrophy of tonsils alone)] is on two lines, one covered and one uncovered (lines 395 STREPTOCOCCAL SORE THROAT AND SCARLET FEVER; VINCENT'S DISEASE; ULCER OF TONSIL; UNILATERAL HYPERTROPHY OF TONSIL, 574 CHRONIC DISEASE OF TONSILS AND ADENOIDS). Similar code 474.10 (Hypertrophy of tonsil with adenoids) only appears on the uncovered line only (line 574). The intention was to cover unilateral tonsillar hypertrophy, which is outlined in GN 36. However, this intent is not clear to health plans and other partners. The ICD-10 equivalent (J35.1) to 474.11 is also on both lines.

From Ms. Andrews:

I have a question about two very similar conditions on the prioritized list. ICD-9 code 474.11 enlarged tonsils alone or tonsillar hypertrophy is a covered code. 474.10 enlarged tonsils and adenoids is below the line. When we make referrals to ENT as PCP, our providers are not able to determine if the adenoids are enlarged as only the tonsils are visualized, and we refer for dx code 474.11 which does not require an EOCCO referral #. ENT typically diagnoses 474.10, and then have to ask for a retroactive referral. Why is there such a difference between these two very similar codes?

GUIDELINE NOTE 36, ADENOTONSILLECTOMY FOR INDICATIONS OTHER THAN OBSTRUCTIVE SLEEP APNEA

Lines 49,84,395,574

Tonsillectomy/adenotonsillectomy is an appropriate treatment for patients with:

- A) Five documented attacks of strep tonsillitis in a year or 3 documented attacks of strep tonsillitis in each of two consecutive years where an attack is considered a positive culture/screen and where an appropriate course of antibiotic therapy has been completed;
- B) Peritonsillar abscess requiring surgical drainage; or,
- C) Unilateral tonsillar hypertrophy in adults; unilateral tonsillar hypertrophy in children with other symptoms suggestive of malignancy.

See Guideline Note 118 for diagnosis and treatment of obstructive sleep apnea in children.

HERC Staff Recommendations:

- 1) Adopt a new guideline note as shown below
 - a) Not appropriate to modify GN36 as the diagnosis code may be used for consultation and non-surgical treatment as well as for tonsillectomy

GUIDELINE NOTE XXX, UNILATERAL TONSILLAR HYPERTROPHY

Lines 395,574

ICD-9 474.11/ICD-10 J35.1 is included on line 395 only for 1) unilateral tonsillar hypertrophy in adults and 2) unilateral tonsillar hypertrophy in children with other symptoms suggestive of malignancy. Bilateral tonsillar hypertrophy and unilateral tonsillar hypertrophy in children without other symptoms suggestive of malignancy are included only on line 574.

Hemangiomas

Question: What changes should be made to the new hemangioma line and guideline?

Question source: DMAP Hearings Division and HERC staff

Issue: A new line for hemangiomas was created as part of the ICD-10 review. This change was made in response to the fact that the hemangioma ICD-9 code most often used was Ancillary, and intended for manual review to see if treatment was appropriate. The HERC decided to define when treatment should be undertaken by putting this code on the Prioritized List with a guideline note. This new line and guideline note will first appear on the January, 2015 Prioritized List.

Recently, the DMAP Hearings Division received a request for coverage for treatment of a hemangioma diagnosis code (ICD-9 228.09) which is still present on the Ancillary List. HERC staff determined that this code was more appropriate for the Prioritized List and the equivalent ICD-10 code was on a line on this list. HERC staff also determined that further definition of when treatment of a hemangioma is NOT covered is required.

January 1, 2015 Prioritized List:

Line: 326

Condition: DERMATOLOGIC HEMANGIOMAS, COMPLICATED (See Guideline Note 13)

Treatment: MEDICAL THERAPY

ICD-9: 228.01

ICD-10: D18.01

CPT: 11300-11446,12031,12032,13100-13151,17106-17108,21011-21014,21552,21554,21931-21933,22901-22903,23071,23073,24071,24073,25071,25073,26111,26113,27043,27045,27337,27339,27632,27634,28039,28041,40500-40530,40810-40816,40820,41116,41826,42104-42107,42160,42808,69145,98966-98969,99051,99060,99070,99078,99201-99215,99281-99285,99341-99355,99358-99378,99381-99404,99408-99412,99429-99449,99487-99496,99605-99607

HCPCS: G0396,G0397,G0463

GUIDELINE NOTE 13, HEMANGIOMAS, COMPLICATED

Line 326

Hemangiomas are covered on this line when they are ulcerated, infected, recurrently hemorrhaging, or function-threatening (e.g. eyelid hemangioma).

ICD-9 Code	Code Description	Current Placement
228.00	Hemangioma of unspecified site	636 BENIGN NEOPLASMS OF SKIN AND OTHER SOFT TISSUES
228.01	Hemangioma of skin and subcutaneous tissue	326 DERMATOLOGIC HEMANGIOMAS, COMPLICATED 636
228.02	Hemangioma of intracranial structures	130 BENIGN NEOPLASM OF THE BRAIN AND SPINAL CORD
228.03	Hemangioma of retina	100 DIABETIC AND OTHER RETINOPATHY
228.04	Hemangioma of intra-abdominal structures	130 BENIGN NEOPLASM OF THE BRAIN AND SPINAL CORD
228.09	Hemangioma of other sites	Ancillary

HERC staff recommendations:

- 1) Add 228.09 (Hemangioma of other sites) to line 636 BENIGN NEOPLASMS OF SKIN AND OTHER SOFT TISSUES
 - a. Advise DMAP to remove 228.09 from the Ancillary list
 - b. Will match the ICD-10 equivalent (D18.09) placement
- 2) Remove 228.04 (Hemangioma of intra-abdominal structures) from line 130 and add to line 647 BENIGN NEOPLASMS OF DIGESTIVE SYSTEM
 - a. Inappropriate placement; used to code for hemangiomas of the liver and GI tract, peritoneum and retroperitoneum
 - b. ICD-10 equivalent code D18.03 is on line 647
- 3) Make the changes shown below to GN13

GUIDELINE NOTE 13, HEMANGIOMAS, COMPLICATED

Line 326, [636](#)

[Dermatologic](#) ~~H~~emangiomas ([ICD-9 228.01](#)) are ~~covered~~ [included](#) on ~~this~~-line [326](#) when they are ulcerated, infected, recurrently hemorrhaging, or function-threatening (e.g. eyelid hemangioma). [Otherwise, they are included on line 636.](#)

Intraocular Steroid Implants

Question: what conditions should be paired with intraocular steroid implants?

Question source: Allergan, Inc.

Issue: Dexamethasone intravitreal implants (brand name Ozurdex) has recently been FDA approved for a new indication: the treatment of diabetic macular edema in patients who are pseudophakic or are phakic and scheduled for cataract surgery. Allergan Inc. is requesting that the HERC re-review coverage of dexamethasone intravitreal implants for this new indication.

Pseudophakic refers to having an artificial lens in place in the eye (i.e. previously having had cataract removal surgery). Phakic and scheduled for cataract surgery refers to patients with cataracts severe enough to require surgery who are planning to undergo cataract removal surgery in the near future. These two groups have been singled out as they do not suffer from one of the major adverse effects and increased costs of intraocular steroid implants, which is the formation of cataracts and the subsequent increased cost of cataract removal/lens implantation surgery. Previous studies have found that intraocular steroid implants are somewhat effective at treatment macular edema in the short term, but lead to long term complications (cataracts, glaucoma) which make the implants less effective and cost-effective than alternative treatments such as systemic steroid therapy.

Current FDA approved indications for dexamethasone intravitreal implants:

Retinal Vein Occlusion

OZURDEX® (dexamethasone intravitreal implant) is a corticosteroid indicated for the treatment of macular edema following branch retinal vein occlusion (BRVO) or central retinal vein occlusion (CRVO).

Posterior Segment Uveitis

OZURDEX® is indicated for the treatment of noninfectious uveitis affecting the posterior segment of the eye.

Diabetic Macular Edema

OZURDEX® is a corticosteroid indicated for the treatment of diabetic macular edema in patients who are pseudophakic or are phakic and scheduled for cataract surgery.

At several VBBS and HERC meetings in 2013, dexamethasone intravitreal implants (Ozurdex) and fluocinolone intraocular implants (brand name Retisert) were reviewed for various indications.

- 1) Uveitis: P&T reviewed this indication in January, 2013 and concluded "There is low quality evidence that there is no difference in visual acuity outcomes between fluocinolone acetonide intravitreal implant and standard of care with systemic corticosteroids for the treatment of noninfectious uveitis. There is also low quality evidence that fluocinolone intravitreal implant may control inflammation in the eye faster and more frequently than standard of care, although both approaches decrease inflammation." These treatments were added to the line 106 DIABETIC AND OTHER RETINOPATHY with a guideline specifying that they appeared on this line for pairing with uveitis which was "chronic, non-infectious, and affecting the posterior segment of the eye, and there has been appropriate trial and failure, or intolerance of therapy, with local and systemic corticosteroids and/or immunosuppressive agents." Line 106 is due to split into 3 lines in January, 2015. The CPT codes will only be included on line 363

CHORIORETINAL INFLAMMATION for the ICD-10 List. HERC staff was also charged with finding more information on the effectiveness of treatment on intermediate uveitis, which is a non-FDA approved indication for implantable steroids.

- 2) Diabetic macular edema: P&T reviewed this indication in January, 2013 and found "There is low quality evidence demonstrating potential benefit of fluocinolone acetonide intravitreal implant and dexamethasone implant for the treatment of diabetic macular edema, however significant complications have also been reported. There is insufficient evidence to support the use for diabetic macular edema or other off-label indications." Intravitreal steroid implants were not added as a treatment for diabetic macular edema.
- 3) Central retinal vein occlusion: P&T reviewed these treatments in January, 2013 and concluded "There is moderate quality evidence demonstrating efficacy of dexamethasone intravitreal implant following central retinal vein occlusion." At the October, 2013 meeting VBBS/HERC added CPT 67028 only to line 445 VENOUS TRIBUTARY (BRANCH) OCCLUSION; CENTRAL RETINAL VEIN OCCLUSION with a guideline specifying coverage was only for those individuals who have failed anti-VEGF therapy.
- 4) Branch retinal vein occlusion: This condition was not included in the P&T review conclusions. It was included in a NICE review that recommended coverage in certain limited situations. At the VBBS/HERC meetings, the commission felt that there was not enough evidence presented to make a decision on coverage, and HERC staff was directed to search for more evidence of effectiveness of intravitreal steroids for this condition.
- 5) Other conditions: the CPT code used for intravitreal steroid injections appears on line 387 CENTRAL SEROUS CHORIORETINOPATHY, but no evidence was found for use for intravitreal steroids for any condition on that line. A coding specification was added to line 387: "CPT 67027 (Implantation of intravitreal drug delivery system) is included on this line for use with medications other than intraocular steroid implants."

Current Prioritized List status:

CPT 67027 (Implantation of intravitreal drug delivery system) and 67028 (Intravitreal injection of a pharmacologic agent) are found on lines

100 DIABETIC AND OTHER RETINOPATHY

284 RETINAL DETACHMENT AND OTHER RETINAL DISORDERS

290 COMPLICATIONS OF A PROCEDURE ALWAYS REQUIRING TREATMENT

323 PURULENT ENDOPHTHALMITIS

387 CENTRAL SEROUS CHORIORETINOPATHY

445 VENOUS TRIBUTARY (BRANCH) OCCLUSION; CENTRAL RETINAL VEIN
OCCLUSION (67028 only)

453 DEGENERATION OF MACULA AND POSTERIOR POLE (67028 only)

**GUIDELINE NOTE 116, INTRAOCULAR STEROID IMPLANTS FOR CHRONIC NON-
INFECTIOUS UVEITIS**

Line 100

Intraocular steroid implants (CPT 67027, 67028) are only included on Line 100 for pairing with uveitis (ICD-9-CM codes 363.0x, 363.1x, 363.20 and 363.22), and only when the following conditions are met: uveitis is chronic, non-infectious, and affecting the posterior segment of the eye, and there has been appropriate trial and failure, or intolerance of therapy, with local and systemic corticosteroids and/or immunosuppressive agents.

**GUIDELINE NOTE 117, INTRAOCULAR STEROID IMPLANTS FOR CENTRAL RETINAL
VEIN OCCLUSION**

Line 445

Intraocular steroid implants (CPT 67028) are only included on Line 445 for treatment of central retinal vein occlusion (ICD-10-CM code H34.81x) in those individuals who have failed anti-VEGF therapy.

Updated evidence review (late 2012-current)

Uveitis

- 1) **Kempen 2011**, MUST trial
 - a. Included in 2013 P&T review
 - b. Included patients with posterior, intermediate, and panuveitis
 - c. Implants and systemic steroids found to have the same impact on visual outcomes; however, implants had much higher rates of cataract and glaucoma development
 - d. Objective visual outcomes for intermediate, posterior, and panuveitis improved over 24 months in both the implant and systemic treatment groups. No breakdown in outcomes was given for these three groups
 - e. Authors also note: "Nearly all phakic eyes receiving implant therapy can be expected to require cataract surgery, and approximately 25% of implant-treated eyes will require glaucoma surgery over 24 months."
- 2) **Comyn 2013**, review of intraocular implants for various eye conditions
 - a. 3 studies reviewed for intraocular steroid implants for uveitis. All studies mixed intermediate, posterior and panuveitis
 - i. 2 studies of fluocinolone
 1. 1 study found lower rate of recurrence with implant vs systemic steroids, while 1 study (MUST above) found no difference between implant and systemic therapy
 2. Very high rates of cataract surgery and interventions for increased intraocular pressure
 - ii. 2 studies of dexamethasone
 1. 1 study showed short term benefit in reducing inflammation (2 months).
 2. 1 retrospective study in children found improvement in vision and ability to spare children systemic steroid therapy

Diabetic macular edema

- 1) **NICE 2013**, evidence review for fluocinolone implant for diabetic macular edema
 - a. Fluocinolone acetonide intravitreal implant is recommended as an option for treating chronic diabetic macular oedema that is insufficiently responsive to available therapies only if the implant is to be used in an eye with an intraocular (pseudophakic) lens
 - b. Decision based on 2 manufacturer sponsored trials (N=956 patients)
 - i. Fluocinolone implant improved best corrected visual acuity significant over sham implant
 - ii. Significantly more cataracts and increased intraocular pressure in the steroid group vs sham
 - c. Economic study determined ICER of £ 22,655 to £ 26,526 per QALY gained for fluocinolone acetonide intravitreal implant compared with optimised standard of care.
 - d. Previous NICE review had found no evidence of benefit of fluocinolone for all diabetic macular edema
 - e. Subgroup of patients with previous cataract surgery/intraocular lens in place were separated for analysis as this group removed one of the outstanding complications/increased cost from the steroid treatment (cataract development and increased cost from cataract surgery)

- 2) **Comyn 2013**, review of intraocular implants for various eye conditions
 - a. 2 trials of fluocinolone implant for diabetic macular edema reviewed
 - i. Both showed short term gains in visual acuity vs sham, but differing results for long term improvement at 3 yrs
 - ii. Very high rate of complications in implant group
 1. 90-91% of patients with implant required cataract surgery!
 2. 8-33% required surgical intervention for increased intraocular pressure
 - b. 1 trial of dexamethasone implant for DME reviewed
- 3) **Qi 2012**, meta-analysis of triamcinolone implant vs subtenon triamcinolone for diabetic macular edema
 - a. N=5 studies (172 eyes)
 - b. No benefit seen from either method at 6 months on visual acuity
 - i. Short term improvement seen over sham, but long term data not yet available

Central/branch retinal vein occlusion

- 4) **Comyn 2013**, review of intraocular implants for various eye conditions
 - a. 4 studies reviewed for CRVO
 - i. 1 study of triamcinolone injection showed it significantly improved visual acuity at 12 months vs grid laser
 - ii. 2 trials of dexamethasone implant showed initial improvement, but no sustained improvement at 12 months vs sham
 - iii. 1 study of fluocinolone implant showed no benefit and 100% of phakic eyes required cataract surgery and 30% of eyes required surgical intervention for increased intraocular pressure
 - b. 1 study for BRVO
 - i. 1 trial of dexamethasone implant showed no long term improvement vs sham, and high rates of cataract surgery in the implant arm
 1. Implants would need to be reinserted every 3-4 months based on trial results and therefore would increase adverse events
- 5) **Pielen 2013**, systematic review of various intravitreal therapies for CRVO/BRVO
 - a. N=11 studies (4 CRVO, 5 BRVO, 2 both)
 - a. 2 studies with triamcinolone
 - b. 1 study with dexamethasone
 - b. CRVO: Triamcinolone stabilized visual acuity for CRVO. Dexamethasone implant had no difference compared to sham
 - c. BRVO: No significant effect was found for triamcinolone at 12 months. BRVO patients treated with dexamethasone implant deteriorated to approximately +6 letters after temporary visual improvement at month 12
 - d. Significantly greater cataract progression with intraocular steroids
 - e. All anti-VEGF agents show a better gain of BCVA than steroids at month 12 (CRVO, BRVO).
 - f. In BRVO best improvement of BCVA was found for ranibizumab at month 12
- 6) **Gawaily 2010**, Cochrane review of intravitreal steroids for CRVO with macular edema vs observation
 - a. Included steroid implants as a possible therapeutic modality
 - b. No RCTS identified that met inclusion criteria (mainly due to lack of observation arm)

- c. Conclusions: There is inadequate evidence for the use of intravitreal steroids for CRVO-ME due to a paucity of RCTs and well-designed observational studies on the topic; therefore, it is still an experimental procedure.
- 7) **NICE 2011**, technology appraisal of dexamethasone implant for treatment of macular edema resulting from central retinal vein occlusion
 - a. Dexamethasone intravitreal implant is recommended as an option for the treatment of macular oedema following central retinal vein occlusion.
 - b. Dexamethasone intravitreal implant is recommended as an option for the treatment of macular oedema following branch retinal vein occlusion when: treatment with laser photocoagulation has not been beneficial, or treatment with laser photocoagulation is not considered suitable because of the extent of macular haemorrhage.
 - c. Data analysis
 - i. Dexamethasone for BRVO showed no difference in visual acuity compared to sham at 180 days, although there was significant improvement at earlier time points
 - ii. Dexamethasone for BRVO was significantly better than sham only for those patients with previous treatment with laser coagulation
 - 1. In the subgroup with BRVO and previous laser therapy 22.2% had an improvement in BCVA from baseline of at least 15 letters at day 30 compared with 2.8% in the sham group ($p = 0.028$). Differences between the dexamethasone and sham groups were also statistically significant at days 60 ($p < 0.001$), 90 ($p = 0.011$) and 180 ($p = 0.022$).

Summary

- 1) *Uveitis*: some benefit seen with steroid implants; however, systemic steroid therapy appears to be equally effective. High rates of complications seen, including cataracts and increased intraocular pressure. Studies did not differentiate between posterior, intermediate, and panuveitis. The treatment appears to have an equal level of effectiveness and side effects for these 3 conditions.
- 2) *Macular edema*: Short-term benefit seen for intraocular steroid implants, but at the cost of an exceedingly high rate of complications (near 100% rate of cataract surgery required, high rates of surgical intervention for increased intraocular pressure). NICE concluded therapy was only cost-effective in eyes which already had undergone cataract surgery to reduce this complication risk/cost.
- 3) *Retinal vein occlusion*: For central retinal vein occlusion, some benefit seen with intraocular steroid implants, but anti-VEGF therapy has better outcomes with fewer complications. For branch retinal vein occlusion, studies have found no long term benefit while significant complications were seen. NICE completed a subgroup analysis and found that patients with BRVO who had previously undergone laser grid therapy had a significant improvement in visual acuity with steroid implants compared to sham.

Pars Planitis

Issue: Pars planitis (ICD-9 363.21/ICD-10 H30.2x) is currently on line 387 CENTRAL SEROUS CHORIORETINOPATHY. However, this diagnosis is a type of uveitis which is more appropriate for another line with other uveitis diagnosis codes and treatments.

The placement of pars planitis was reviewed in October, 2009. At that time, pars planitis was on line 672, SENSORY ORGAN CONDITIONS WITH NO OR MINIMALLY EFFECTIVE TREATMENTS OR NO TREATMENT NECESSARY. Several community ophthalmologists asked the HSC to move this diagnosis from its very low, uncovered line, to a higher, covered line. This request was denied as a review found little need for treatment.

In June, 2010, as part of the biennial review at that time, several ophthalmologists again requested that the placement of pars planitis, among several other diagnoses, be reviewed. The ophthalmologists provided evidence that there were effective treatments for pars planitis, and that treatment was necessary in many cases. During this review, a new line was created, entitled Central Serous Retinopathy. This line was originally supposed to contain a variety of unrelated ophthalmologic diagnoses which only required treatment in some cases. Central serous retinopathy, pars planitis, and hypotony were placed on this line with a guideline regarding when certain of these conditions, including pars planitis, should be treated. Initially, this line was scored to an unfunded position, but it was later moved into the funded region.

On review, pars planitis is also known as intermediate uveitis or posterior cyclitis. Its treatments are much the same as posterior uveitis and panuveitis. HERC staff recommends moving from the central serous retinopathy line to the very similar priority chorioretinal inflammation line.

Current guideline note:

GUIDELINE NOTE 10, CENTRAL SEROUS CHORIORETINOPATHY AND POSTERIOR CYCLITIS

Line 387

Central serous chorioretinopathy (H35.71x) is included on this line only for treatment when the condition has been present for 3 months or longer. Posterior Cyclitis (H30.2x) should only be treated in patients with 20/40 or worse vision.

HERC staff recommendations:

- 1) Add panuveitis (ICD-9 360.12/ICD-10 H44.11x) and sympathetic uveitis (ICD-9 360.11/ICD-10 H44.131-9) to line 363 CHORIORETINAL INFLAMMATION and remove from line 269 ACUTE, SUBACUTE, CHRONIC AND OTHER TYPES OF IRIDOCYCLITIS
 - a. Panuveitis and sympathetic uveitis were added to this line when it was created during the ICD-10 ophthalmology review but were not moved from line 269 in apparent mistake
- 2) Add pars planitis (ICD-9 363.21/ICD-10 H30.2x) to line 363 CHORIORETINAL INFLAMMATION and remove from line 387 CENTRAL SEROUS CHORIORETINOPATHY
- 3) Modify GN 10 as shown below
 - a. Add new line number for pars planitis
- 4) Modify GN 116 as shown below
 - a. Add steroid implants as treatment for intermediate uveitis and panuveitis
 - i. Add diagnostic code for panuveitis and pars planitis to the guideline
 - ii. Remove wording restricting coverage to posterior disease
 - b. Evidence is similar for posterior, intermediate and panuveitis
 - i. Evidence of effectiveness is poor for all conditions and treatment should be a third line therapy
 - c. Note: FDA approval is only for posterior uveitis
- 5) Modify GN 177 as shown below
 - a. Change current guideline regarding coverage of intraocular steroids for central retinal vein occlusion
 - i. Evidence supporting efficacy is for treatment of the resulting macular edema
 - b. Add coverage for intraocular steroids for branch retinal vein occlusion only for patients who have failed laser photocoagulation
 - i. Follows NICE guidance and evidence
- 6) Add intraocular steroid implants as a treatment for diabetic macular edema in patients with artificial lens or with planned cataract removal surgery
 - a. Add CPT 67027 and 67028 to line 100 DIABETIC AND OTHER RETINOPATHY
 - b. Add the new guideline shown below to line 100
 - c. Alternative: continue with current non-coverage due to exceedingly high rate of complications

GUIDELINE NOTE 10, CENTRAL SEROUS CHORIORETINOPATHY AND POSTERIOR CYCLITIS

Line [363](#), 387

Central serous chorioretinopathy (H35.71x) is included on ~~this~~ line [387](#) only for treatment when the condition has been present for 3 months or longer. Posterior Cyclitis (H30.2x) should only be treated in patients with 20/40 or worse vision.

GUIDELINE NOTE 116, INTRAOCULAR STEROID IMPLANTS FOR CHRONIC NON-INFECTIOUS UVEITIS

Line ~~100~~363

Intraocular steroid implants (CPT 67027, 67028) are only included on Line ~~100~~363 for pairing with uveitis (ICD-9-CM codes 360.12, 363.0x, 363.1x, 363.2x, ~~363.20 and 363.22~~/ICD-10-CM codes H30.0xx, H30.1xx, H30.89x, H30.9xx, H44.11x), and only when the following conditions are met: uveitis is chronic, non-infectious, ~~and affecting the posterior segment of the eye~~, and there has been appropriate trial and failure, or intolerance of therapy, with local and systemic corticosteroids and/or immunosuppressive agents.

GUIDELINE NOTE 117, INTRAOCULAR STEROID IMPLANTS FOR CENTRAL RETINAL VEIN OCCLUSION

Line 445

Intraocular steroid implants (CPT 67028) are only included on Line 445 for treatment of macular edema due to

- 1) central retinal vein occlusion (ICD-9-CM 362.35/ICD-10-CM code H34.81x) in those individuals who have failed anti-VEGF therapy.
- 2) branch retinal vein occlusion (ICD-9-CM 362.36/ICD-10-CM code H34.83x) when treatment with laser photocoagulation has not been beneficial, or treatment with laser photocoagulation is not considered suitable because of the extent of macular hemorrhage.

GUIDELINE NOTE XXX_INTRAOCULAR STEROID IMPLANTS FOR DIABETIC MACULAR EDEMA

Line 100

Intraocular steroid implants (CPT 67027, 67028) is included on this line for treating chronic diabetic macular edema (ICD-9 362.07/ ICD-10 E11.311) only when

- 1) there has been insufficient response to available therapies, and
- 2) the implant is to be used in an eye with an intraocular (pseudophakic) lens or in a phakic eye scheduled for cataract surgery.

Randomized Comparison of Systemic Anti-inflammatory Therapy Versus Fluocinolone Acetonide Implant for Intermediate, Posterior, and Panuveitis: The Multicenter Uveitis Steroid Treatment Trial

The Multicenter Uveitis Steroid Treatment (MUST) Trial Research Group*

*Writing Committee: John H. Kempen, MD, PhD,^{1,2,3} Michael M. Altaweel, MD,⁴ Janet T. Holbrook, PhD, MPH,^{5,6} Douglas A. Jabs, MD, MBA,^{5,6,8,9} Thomas A. Louis, PhD,⁷ Elizabeth A. Sugar, PhD,^{5,6,7} Jennifer E. Thorne, MD, PhD^{5,6,10}

Objective: To compare the relative effectiveness of systemic corticosteroids plus immunosuppression when indicated (systemic therapy) versus fluocinolone acetonide implant (implant therapy) for noninfectious intermediate, posterior, or panuveitis (uveitis).

Design: Randomized controlled parallel superiority trial.

Participants: Patients with active or recently active uveitis.

Methods: Participants were randomized (allocation ratio 1:1) to systemic or implant therapy at 23 centers (3 countries). Implant-assigned participants with bilateral uveitis were assigned to have each eye that warranted study treatment implanted. Treatment-outcome associations were analyzed by assigned treatment for all eyes with uveitis.

Main Outcome Measures: Masked examiners measured the primary outcome: change in best-corrected visual acuity from baseline. Secondary outcomes included patient-reported quality of life, ophthalmologist-graded uveitis activity, and local and systemic complications of uveitis or therapy. Reading Center graders and glaucoma specialists assessing ocular complications were masked. Participants, ophthalmologists, and coordinators were unmasked.

Results: On evaluation of changes from baseline to 24 months among 255 patients randomized to implant and systemic therapy (479 eyes with uveitis), the implant and systemic therapy groups had an improvement in visual acuity of +6.0 and +3.2 letters ($P = 0.16$, 95% confidence interval on difference in improvement between groups, -1.2 to $+6.7$ letters, positive values favoring implant), an improvement in vision-related quality of life of +11.4 and +6.8 units ($P = 0.043$), a change in EuroQol-EQ5D health utility of +0.02 and -0.02 ($P = 0.060$), and residual active uveitis in 12% and 29% ($P=0.001$), respectively. Over the 24 month period, implant-assigned eyes had a higher risk of cataract surgery (80%, hazard ratio [HR] = 3.3, $P < 0.0001$), treatment for elevated intraocular pressure (61%, HR=4.2, $P < 0.0001$), and glaucoma (17%, HR=4.2, $P = 0.0008$). Patients assigned to systemic therapy had more prescription-requiring infections than patients assigned to implant therapy (0.60 vs 0.36/person-year, $P=0.034$), without notable long-term consequences; systemic adverse outcomes otherwise were unusual in both groups, with minimal differences between groups.

Conclusions: In each treatment group, mean visual acuity improved over 24 months, with neither approach superior to a degree detectable with the study's power. Therefore, the specific advantages and disadvantages identified should dictate selection between the alternative treatments in consideration of individual patients' particular circumstances. Systemic therapy with aggressive use of corticosteroid-sparing immunosuppression was well tolerated, suggesting that this approach is reasonably safe for local and systemic inflammatory disorders.

Financial Disclosure(s): Proprietary or commercial disclosure may be found after the references. *Ophthalmology* 2011;118:1916–1926 © 2011 by the American Academy of Ophthalmology.



*Group members listed online in Appendix 1 (available at <http://aaojournal.org>).

Uveitis (intraocular inflammation) is an important cause of visual impairment.^{1,2} Intermediate, posterior, and panuveitis^{3,4} are the forms of uveitis most likely to cause vision

loss.^{5,6} Because of its earlier onset, uveitis results in a longer duration of blindness and more economic cost per case than more common age-related ocular diseases.⁷



Corticosteroid intravitreal implants vs. ranibizumab for the treatment of vitreoretinal disease

Oliver Comyn^{a,c,d}, Sue L. Lightman^{c,d}, and Philip G. Hykin^{b,c}

Purpose of review

Three long-acting corticosteroid implants are now available for the treatment of retinal disease, offering control of macular edema and inflammation for between 6 months and up to 3 years. This review evaluates their efficacy and side-effect profile in comparison with the antivascular endothelial growth factor agent ranibizumab in diabetic macular edema, retinal vein occlusion, pseudophakic macular edema, and uveitis.

Recent findings

Trials of ranibizumab in diabetic macular edema have demonstrated excellent efficacy without serious safety concerns to date. Fluocinolone acetonide implants can be considered, but have a high risk of cataract and sequelae from intraocular pressure rise. In retinal vein occlusion, both ranibizumab and Ozurdex have been shown to be effective, although their relative efficacy has not been determined in head-to-head clinical trials. In pseudophakic and uveitic macular edema, steroid implants are probably the first choice therapy, although there is evidence that ranibizumab is effective. For choroidal neovascularization secondary to inflammatory disease, ranibizumab is indicated, whereas Retisert has been shown to reduce the risk of uveitis relapse.

Summary

In diabetic macular edema, ranibizumab has shown greater efficacy with fewer side-effects than steroid implants. Both ranibizumab and steroid implants can be considered in retinal vein occlusion, but trials are awaited to determine their relative efficacy.

Keywords

diabetic macular edema, retinal vein occlusion, uveitis

INTRODUCTION

Recent developments in biodegradable polymers coupled with expertise in the delivery of intravitreal drugs have led to the clinical use of several devices releasing corticosteroids into the vitreous for prolonged periods. Offering the advantage of continuous therapy for 6 months to 3 years as opposed to repeated re-injection of a short-acting drug in liquid form, these implants are an exciting addition to the armoury of the retinal specialist.

The potential side-effects of long-acting steroid preparations are cataract and glaucoma, which are more likely to occur in a significant number of patients if the duration of action exceeds 6 months. This may result in additional intraocular surgery in such instances and can be considered a drawback. Their adverse effect profile, together with limited efficacy in retinal vascular disease, has led to their moderate use to date.

In this review, we compare the efficacy and safety profile of ranibizumab (Lucentis; Genentech;

South San Francisco, California, USA), the first monoclonal antibody fragment directed against all isoforms of vascular endothelial growth factor (VEGF) to be licensed for intraocular use, with the three main commercially available slow-release corticosteroid implants: Retisert, Iluvien, and Ozurdex.

Two of the available devices release fluocinolone acetonide, whose low solubility makes it suitable for slow-release without the need for a polymer system. Retisert (Bausch & Lomb, Rochester, New York, USA) is a sterile, nonbiodegradable intravitreal implant,

^aNIHR Biomedical Research Centre, ^bMedical Retina Department, ^cMoorfields Eye Hospital NHS Foundation Trust and ^dUCL Institute of Ophthalmology, London, UK

Correspondence to Mr Philip G. Hykin, FRCOphth, Moorfields Eye Hospital, City Road, London EC1V 2PD, UK. Tel: +44 207 253 3411; e-mail: philhykin@aol.com

Curr Opin Ophthalmol 2013, 24:248–254

DOI:10.1097/ICU.0b013e32835fab27

Fluocinolone acetonide intravitreal implant for treating chronic diabetic macular oedema after an inadequate response to prior therapy (rapid review of technology appraisal guidance 271)

Issued: November 2013

NICE technology appraisal guidance 301

guidance.nice.org.uk/ta301

Contents

1 Guidance	3
2 The technology	4
3 The manufacturer's submission.....	5
ERG comments on the manufacturer's submission	13
Exploratory analyses by the ERG	14
Rapid review of NICE technology appraisal guidance 271: patient access scheme.....	17
Comments on the manufacturer's rapid review submission by the Evidence Review Group	20
4 Consideration of the evidence	22
Clinical effectiveness.....	24
Cost effectiveness	26
Summary of Appraisal Committee's key conclusions.....	33
5 Implementation	40
6 Recommendations for further research	41
7 Related NICE guidance	42
Published.....	42
8 Review of guidance	43
9 Appraisal Committee members, guideline representatives and NICE project team.....	44
9.1 Appraisal Committee members	44
9.2 NICE project team	47
10 Sources of evidence considered by the Committee	48
About this guidance	51

1 Guidance

This guidance replaces NICE technology appraisal guidance 271 issued in January 2013. The review of fluocinolone acetonide intravitreal implant for treatment of chronic diabetic macular oedema after an inadequate response to prior therapy has resulted in a change in the guidance. See [About this guidance](#) for more information.

- 1.1 Fluocinolone acetonide intravitreal implant is recommended as an option for treating chronic diabetic macular oedema that is insufficiently responsive to available therapies only if:
- the implant is to be used in an eye with an intraocular (pseudophakic) lens **and**
 - the manufacturer provides fluocinolone acetonide intravitreal implant with the discount agreed in the patient access scheme.

2 The technology

- 2.1 Fluocinolone acetonide intravitreal implant (Iluvien, Alimera Sciences) contains a corticosteroid that has anti-inflammatory and anti-vascular endothelial growth factor (anti-VEGF) properties. It is administered by intravitreal injection. Each implant contains 190 micrograms of fluocinolone acetonide, releasing 0.2 micrograms/day for approximately 36 months. Fluocinolone acetonide intravitreal implant has a marketing authorisation for 'the treatment of vision impairment associated with chronic diabetic macular oedema considered insufficiently responsive to available therapies'. The summary of product characteristics states that administration in both eyes concurrently is not recommended.
- 2.2 The summary of product characteristics lists the following adverse reactions for fluocinolone acetonide intravitreal implant: cataract, increased intraocular pressure, floaters (myodesopsia), retinal detachments, vitreous haemorrhages or detachments, glaucoma and endophthalmitis. For full details of adverse reactions and contraindications, see the summary of product characteristics.
- 2.3 Fluocinolone acetonide intravitreal implant is available in a 190-microgram implant at a price of £5500 (excluding VAT; 'British National Formulary' [BNF] 65th edition). Costs may vary in different settings because of negotiated procurement discounts.
- 2.4 The manufacturer of fluocinolone acetonide intravitreal implant has agreed a patient access scheme with the Department of Health in which fluocinolone acetonide intravitreal implant will be available with a discount (see [section 5.3](#)). The details of the scheme were provided as commercial in confidence. The Department of Health considered that this patient access scheme does not constitute an excessive administrative burden on the NHS.

3 The manufacturer's submission

The Appraisal Committee ([section 9](#)) considered evidence submitted by the manufacturer of fluocinolone acetonide intravitreal implant and a review of this submission by the Evidence Review Group (ERG; [section 10](#)).

- 3.1 The manufacturer submitted evidence on the clinical and cost effectiveness of fluocinolone acetonide intravitreal implant compared with optimised standard of care and laser photocoagulation monotherapy. The manufacturer did not provide any specific analyses comparing fluocinolone acetonide intravitreal implant with triamcinolone alone or the anti-vascular endothelial growth factor (anti-VEGF) treatments bevacizumab and ranibizumab alone.
- 3.2 The main source of evidence in the manufacturer's submission was a preplanned analysis of data from the FAME A and B randomised controlled trials that evaluated the safety and efficacy of fluocinolone acetonide intravitreal implant for treating diabetic macular oedema. The preplanned analysis focused on duration of diabetic macular oedema, analysing patients who had had the condition for durations above and below the median separately. When the trial was unblinded the median duration was determined to be 3 years. The subgroup in the submission was patients with duration of diabetic macular oedema over 3 years (the manufacturer calculated the duration of diabetic macular oedema as the year of randomisation to treatment minus the year of diagnosis of the disease plus 1).
- 3.3 FAME A and B were 2 identical, randomised, double-blinded, sham injection-controlled multicentre trials conducted over 36 months. The results of the trials were combined and presented in the submission as a single analysis. Patients were randomised 1:2:2 to sham injection, 0.2 micrograms/day (low-dose) or 0.5 micrograms/day (high-dose) fluocinolone acetonide intravitreal implant. Participants in the trials were adults with diabetic macular oedema who were aged between 18 and 85, who had received at least 1 previous laser treatment, whose best corrected visual acuity (BCVA) was ≥ 19 to ≤ 68 letters on the Early Treatment of Diabetic Retinopathy Study (ETDRS) eye chart, and whose central retinal thickness was 250 microns or more at baseline. Exclusion criteria were intraocular pressure over 21 mmHg, and

systolic blood pressure over 180 mmHg or diastolic blood pressure over 105 mmHg.

- 3.4 In both groups (sham injection and low- or high-dose implant) additional treatment with laser photocoagulation was given as needed after week 6. Approximately 61% of the sham injection group and 41% of the fluocinolone acetonide intravitreal implant group received treatment with laser photocoagulation during the study. The mean number of laser treatments was 1.4 and 0.8 in the sham injection and fluocinolone acetonide intravitreal implant groups respectively. Re-treatment with fluocinolone acetonide intravitreal implant was offered at any time after the month 12 assessments if patients experienced vision loss (5 letters or more) or retinal thickening of 50 microns or more compared with their best status in the previous 12 months.
- 3.5 Patients in both groups also received a range of therapies not allowed in the study protocol. These included intravitreal steroids (triamcinolone and dexamethasone), anti-VEGF therapy, vitrectomies and posterior sub-Tenon steroids. The number of off-protocol treatments was higher in the sham injection group than in the fluocinolone acetonide intravitreal implant group (117 compared with 48); approximately 35% of patients in the sham injection group and 13% of patients in the fluocinolone acetonide intravitreal implant group received at least 1 off-protocol treatment. Data from these patients were included in the analysis population.
- 3.6 The primary outcome reported in the FAME trials was the proportion of people with an improvement of 15 or more letters from baseline BCVA at month 24. Secondary outcomes included:
- mean change in BCVA
 - mean change in excess retinal thickness
 - percentage with 3-step (15 letters or more) worsening of ETDRS
 - percentage needing laser photocoagulation.
- 3.7 There were 956 patients enrolled in the FAME trials. Of these, 536 formed the subgroup of patients with chronic diabetic macular oedema for longer than

3 years. Of this subgroup of patients comprehensive data were presented by the manufacturer for the 0.2 micrograms/day implant group only, because only the low dose has been licensed. The resulting number of patients with chronic diabetic macular oedema in the subgroup that formed the basis of the manufacturer's submission was 321 (209 in the 0.2 micrograms/day implant group and 112 in the sham injection group).

- 3.8 The mean age of the patients was 62.9 years in the sham injection group (n=112) and 63.7 years in the 0.2 micrograms/day implant group (n=209). In the trials, the majority of patients had chronic diabetic macular oedema in both eyes at baseline, but only 1 eye was treated. In most cases, the worse-seeing eye was treated.
- 3.9 At month 24, the proportions of patients with chronic diabetic macular oedema who had a ≥ 15 letter increase from baseline BCVA were 13.4% and 34.4% in the sham injection and 0.2 micrograms/day implant groups respectively ($p < 0.001$). The proportions of patients who had a ≥ 15 letter improvement in BCVA at month 36 (13.4% and 34.0% respectively [$p < 0.001$]) were comparable to those observed in month 24. There were numerical increases in mean change in BCVA from baseline in the 0.2 micrograms/day implant group compared with the sham injection group at all evaluations (12 through to 36 months); this was statistically significant at months 30 and 36. At month 36, there was a mean improvement of 7.6 letters in the 0.2 micrograms/day implant treatment group compared with 1.8 letters in the sham injection group ($p < 0.004$).
- 3.10 The FAME trials included an assessment of health-related quality of life using the Visual Function Questionnaire-25 (VFQ-25) at baseline and months 24 and 36. These data were provided in the clinical study reports of the FAME trials. The manufacturer stated that the VFQ-25 was not used in the economic model because it measures overall visual function which is driven by vision in the better-seeing eye, whereas in the FAME trials, the majority of patients had their worse-seeing eye treated. The VFQ-25 values are marked by the manufacturer as academic in confidence and therefore not presented here.

- 3.11 The manufacturer also included laser photocoagulation monotherapy as a relevant comparator for fluocinolone acetonide intravitreal implant. The manufacturer conducted a literature search and identified 1 relevant study: DRCR Protocol B (2008). This study was a phase 3, multicentre, randomised clinical trial conducted in the USA to compare intravitreal triamcinolone with focal/grid laser photocoagulation in patients with diabetic macular oedema. The manufacturer noted that the severity of diabetic macular oedema in the DRCR study was not as great as in the FAME trials. Of the DRCR population, approximately 40% of people had not had their disease treated with laser and there was no stipulation on duration of diabetic macular oedema at randomisation. A comparison of outcomes data as reported in the DRCR and FAME clinical trials was presented. The manufacturer did not use any statistical methods to compare the data indirectly. The manufacturer noted that the proportion of people with a ≥ 15 letter improvement in BCVA in the FAME trials at month 36 demonstrated a numerical difference in favour of fluocinolone acetonide intravitreal implant (34.0% compared with 18% for laser photocoagulation at 24 months in the DRCR Protocol B study).
- 3.12 The manufacturer's submission included data from the FAME trials for ocular adverse events in people with chronic diabetic macular oedema (duration of 3 years or longer). The data suggested that fluocinolone acetonide intravitreal implant is associated with the formation or progression of cataract and increased intraocular pressure. In the FAME trials, 34.4% (72/209) of patients in the fluocinolone acetonide intravitreal implant group experienced increased intraocular pressure or ocular hypertension compared with 14.3% (16/112) in the sham injection group. At baseline, 58.9% (66/112) of people in the sham injection group and 54.5% (114/209) in the fluocinolone intravitreal implant group were phakic (still had their natural lens). Of those who were phakic at baseline, 77.9% of the fluocinolone acetonide intravitreal implant group and 77.0% of the sham injection group had a pre-existing cataract. Cataract surgery was needed by 85.1% (97/114) of the fluocinolone acetonide intravitreal implant group and 36.4% (24/66) of the sham injection group who were phakic at baseline.
- 3.13 The manufacturer provided a subgroup analysis of people with chronic diabetic macular oedema who had treatment in an eye with a pseudophakic lens (that

is, they had already had an operation for cataract removal and had been fitted with an intraocular lens to replace the natural crystalline lens) at entry into the FAME trials. The manufacturer considered the overall treatment effect in this subgroup to be similar to that for the phakic population; however, increased benefits arose through the removal of a known adverse event (advancement of cataract development) and removal of the costs associated with cataract surgery. Of the people with chronic diabetic macular oedema, there were 46 (41.1%) who had a pseudophakic lens in the sham injection group and 95 (45.5%) in the 0.2 micrograms/day fluocinolone-treated group. For this subgroup, at 36 months, 31.6% in the fluocinolone-treated group and 17.4% in the sham injection group had a greater than 15-letter increase from baseline BCVA, giving a difference of 14.2% (95% CI 28.6% to -0.2%).

- 3.14 The economic evidence provided by the manufacturer in its submission comprised a literature review (which identified no relevant published cost-effectiveness studies) and a de-novo cost–utility analysis. The manufacturer's economic evaluation compared fluocinolone acetonide intravitreal implant with the comparator ('optimised standard of care') in the FAME trials for a cohort of patients with chronic diabetic macular oedema. The model also included a comparison of fluocinolone acetonide intravitreal implant with laser photocoagulation using data for laser photocoagulation from the DRCR Protocol B study.
- 3.15 The manufacturer's model included 14 health states (13 BCVA health states and death) which were defined by bands of 5 ETDRS letters in the treated eye. Utility values associated with the 13 BCVA-related health states captured the effect of varying degrees of visual gain or loss on patients' quality of life. The model structure made no distinction between treatment of the better-seeing eye and the worse-seeing eye. The model had a 15-year time horizon and a quarterly cycle length, with costs and benefits both being discounted at 3.5%. For the first 3 years, the distribution of patients across health states was drawn directly from the FAME trials data. Beyond 3 years, a Markov model structure was adopted with transition probabilities being applied.
- 3.16 The base-case analysis assumed that 1 fluocinolone acetonide intravitreal implant is needed every 3 years. Patients needed to have gained 5 or more

ETDRS letters of visual acuity between baseline and month 36 to receive a further implant at month 36. In addition, using data from the FAME trials the manufacturer applied a drop-out rate (for those who withdrew consent, were lost to follow-up or died) to the patients receiving fluocinolone acetonide intravitreal implant who were re-treated at the end of the first 36 months (these details are marked by the manufacturer as commercial in confidence and therefore not presented here). This adjusted re-treatment rate was applied equally to each health state. Patients in both the fluocinolone group and the optimised standard of care group also received laser treatments based on rates from the FAME trials, and in the first 3 years could receive other therapies including triamcinolone, ranibizumab, bevacizumab and dexamethasone, again based on rates in the FAME trials.

- 3.17 In the base-case analysis, it was assumed that 35% of patients would receive bilateral treatment in the optimised standard of care group. The manufacturer assumed that bilateral treatment would be contraindicated for patients treated in the first eye with fluocinolone acetonide intravitreal implant who had a subsequent rise in intraocular pressure greater than 30 mmHg. Therefore the bilateral treatment rate in the fluocinolone group was reduced based on the proportion of patients with raised intraocular pressure observed in the FAME trials.
- 3.18 The clinical efficacy data from the FAME trials were used directly to calculate the number of patients in each of the model health states in each quarter for the first 3 years. Changes after 3 years were extrapolated from the FAME trials data. The FAME data were divided into patients whose disease had responded to and not responded to treatment based on the ETDRS 5-letter criteria. Data were then analysed to determine the numbers of patients whose vision improved or worsened by at least 5 letters each quarter. The average net changes in the last 4 quarters of the FAME trials were used in the model to extrapolate improvements in vision beyond 3 years. For patients receiving another implant at 36 months, the model assumed that 5% of patients in each health state would improve by 5 letters every quarter. For patients not receiving a further implant at 36 months, the model assumed that 3% of patients in each health state would experience a transition to a lower health state every quarter. In the optimised standard of care group and laser group, the model assumed

that 3% of patients in each health state would have a worsening in vision of 5 letters and therefore move to a worse health state every quarter.

- 3.19 The manufacturer did not consider it appropriate to include the VFQ-25 values in the economic model because VFQ-25 is driven by vision in the better-seeing eye, whereas in the FAME trials, the majority of patients had their worse-seeing eye treated (these values are marked by the manufacturer as academic in confidence and are therefore not presented here). Furthermore, the manufacturer stated that a mapping exercise was not considered because there was not a universally accepted mapping process to convert VFQ-25 data to utility scores.
- 3.20 The manufacturer conducted a systematic review to identify utility values reported in the literature for populations with visual impairment. The review of the articles included diabetic macular oedema and other disorders affecting visual acuity (such as age-related macular degeneration). Based on the data available, the manufacturer chose to use time trade-off data from Brown et al. (2000) as the source of utility values for its submission. Brown et al. (2000) was a US study that measured utility values in 5 groups according to visual acuity in the better-seeing eye in a population of patients with age-related macular degeneration. The values estimated by Brown et al. (2000) and the values used within the model ranged from 0.40 in the lowest health state (<20 ETDRS letters) to 0.89 in the highest health state (≥75 ETDRS letters). The Brown et al. (2000) study did not report utility weights in patients with BCVAs between 35 and 50, and therefore the unweighted averages of the utility weights above and below this range were assumed.
- 3.21 For patients who received treatment in both eyes, a 25% bilateral treatment quality-adjusted life year (QALY) uplift was also applied to the aggregate QALYs.
- 3.22 The model did not consider utility decrements due to adverse events, or procedures and interventions for the adverse events. The manufacturer stated that because the utility was calculated for BCVA values, and the BCVA values were based on the trial data, the impact on patient vision of adverse events such as cataract formation was reflected in the BCVA of the treated eye.

- 3.23 The model included the costs of fluocinolone acetonide intravitreal implant as well as laser and other therapies at the rates observed in the FAME trials. Adverse event costs were also included. The manufacturer applied an annual cost of blindness of £6298 to the proportion of patients whose treated-eye BCVA fell below 35 letters.
- 3.24 In its deterministic base case, based on an incremental cost of £11,330 and an incremental QALY value of 0.500, the manufacturer estimated an incremental cost-effectiveness ratio (ICER) without the patient access scheme of £22,655 per QALY gained for fluocinolone acetonide intravitreal implant compared with optimised standard of care.
- 3.25 Following a request for clarification from the ERG the manufacturer provided a revised analysis. The manufacturer acknowledged that the health-related quality of life values from Brown et al. (2000) may not apply to patients having their worse-seeing eye treated. The manufacturer therefore used revised health-related quality of life values to reflect a weighted average of values for people having their worse- and better-seeing eyes treated, taken from a study by Heintz et al. (2012).
- 3.26 As well as changing health-related quality of life values, the manufacturer's revised analysis also amended: male and female mortality rates to revise the pooled annual all-cause mortality risk, the proportion of patients needing bilateral treatment, the percentage of patients needing bilateral treatment and for whom a second fluocinolone acetonide intravitreal implant treatment was not contraindicated, the quality of life uplift from bilateral treatment to 10%, and the unadjusted response rate in the fluocinolone group to a rate based upon a 10-letter re-treatment criterion.
- 3.27 The manufacturer's amendments reduced the estimate of cost effectiveness for fluocinolone acetonide intravitreal implant compared with optimised standard of care from £22,655 to £19,268 per QALY gained without the patient access scheme, based on an incremental cost of £11,927 and an incremental QALY value of 0.619.

ERG comments on the manufacturer's submission

- 3.28 The ERG commented that the 3-year data used to inform the first 3 years of the economic model were robust, although the more usual modelling approach would have been to use transition probability matrices. The structure of the model means it cannot be manipulated during this 3-year period to explore different scenarios.
- 3.29 The ERG commented that it would have been more appropriate for the manufacturer to use a model structure that modelled patients as having 2 eyes, rather than undertaking an ad hoc adjustment to the output of a model in which patients only had 1 eye. The ERG noted that the FAME trials had a reasonable proportion of patients who had their better-seeing eye treated, and that the rate of chronic diabetic macular oedema in the other eye was high.
- 3.30 The ERG also noted that the distribution between health states for patients whose disease had responded to and not responded to fluocinolone acetonide intravitreal implant at 36 months was modelled as being a constant percentage of the overall patient distribution at 36 months. The ERG commented that this approach was not justified, and could lead to bias in the estimates of cost effectiveness for fluocinolone acetonide intravitreal implant.
- 3.31 The ERG noted that the manufacturer's base-case model applied a re-treatment criterion of a minimum 5-letter improvement between baseline and 36 months. The ERG commented that a more realistic criterion might be a minimum 10-letter improvement between baseline and 36 months, which the manufacturer applied in response to the clarification request from the ERG. However, the ERG commented that this only changed the number of patients not receiving another fluocinolone acetonide intravitreal implant at 36 months and did not affect their distribution across health states.
- 3.32 In the manufacturer's submission clinical effectiveness beyond 3 years was extrapolated from the FAME trials data. The ERG commented that rather than including the proportions of patients whose disease improved or worsened each quarter of a year, the proportions were netted. The ERG considered that the reasons for analysing the data in this way were unclear.

- 3.33 The ERG commented that because in the FAME trials patients had only 1 eye treated, assumptions were needed about rates of bilateral treatment. In the base-case analysis, it was assumed that 35% of patients would need bilateral treatment; this percentage was increased in the revised analysis (these details are marked by the manufacturer as academic in confidence and therefore not presented here). The ERG considered that the manufacturer's revision was too high because a proportion of patients would not have visual impairment in both eyes because of diabetic macular oedema, and a proportion would not be able to have both eyes treated because of raised intraocular pressure or other reasons.
- 3.34 The ERG commented that there was considerable uncertainty about the appropriateness of the utility values used in the model. The original submission used utility values that related to sight in the better-seeing eye. The ERG considered there were limitations in the data (based on Heintz et al. 2012) used in the revised analysis because these provided only 3 quality of life values over 13 health states, were based on small patient numbers, and were non-monotonic.

Exploratory analyses by the ERG

- 3.35 Initially, the ERG made a series of revisions to address what it considered to be possible errors in the model. These were:
- a change to the formulae for averaging mortality between male and female rates
 - a change to the formulae for applying the yearly natural discontinuation rate (the detail of which is commercial in confidence) in the cohort flow cells
 - a change to the formula for the percentage of patients remaining on fluocinolone acetonide intravitreal implant after year 9.

The cumulative impact of correcting the 3 errors was to increase the ICER without the patient access scheme from £22,655 to £26,526 per QALY gained for fluocinolone acetonide intravitreal implant compared with optimised standard of care.

3.36 The ERG also conducted additional exploratory analyses changing:

- how the cost of fluorescein angiography needed before each laser administration was applied in the model
- the number of laser administrations per patient
- the unit costs for adverse event procedures
- how the proportions of people whose disease improved and worsened each quarter year after 36 months was calculated
- the rate of bilateral treatment
- the cost of blindness
- the cost and quality of life uplift applied for bilateral treatment for patients in the fluocinolone group whose condition needed bilateral treatment but for whom a second implant was contraindicated.

The cumulative impact of these changes resulted in ICERs of £37,740 for the 5-letter response criterion (with an incremental cost of £14,569 and an incremental QALY of 0.386) and £35,940 for the 10-letter response criterion (with an incremental cost of £12,736 and an incremental QALY of 0.354) without the patient access scheme.

3.37 The ERG also conducted further exploratory sensitivity analyses with a particular focus on the source of utility values. These additional analyses modelled the impact of changes in the BCVA of the better-seeing eye using Brown et al. (1999). This US study of 325 patients measured utility values in 12 groups according to visual acuity in the better-seeing eye in a population of patients with impaired vision in at least 1 eye. The values estimated by Brown et al. (1999) used within the model vary from 0.54 in the lowest health state (<20 ETDRS letters) to 0.89 in the highest health state (≥75 ETDRS letters), giving a range of 0.350 across the health states of the model. The ERG commented that if its cost-effectiveness estimates without the patient access scheme for fluocinolone acetonide intravitreal implant compared with optimised standard of care were used as a starting point (£37,740 per QALY gained for the 5-letter response criterion and £35,940 per QALY gained for the 10-letter

response criterion), applying the health-related quality of life values taken from Brown et al. (1999) would suggest cost-effectiveness estimates of £66,744 per QALY gained (5-letter response criterion) and £64,249 per QALY gained (10-letter response criterion).

3.38 The ERG explored the effect of using health-related quality of life values from a regression analysis in the manufacturer's submission for [Ranibizumab for the treatment of diabetic macular oedema](#) (NICE technology appraisal guidance 237, replaced by [technology appraisal guidance 274](#)), and from Brown et al. (1999) and Brown et al. (2000). The ERG noted that there was uncertainty around the health-related quality of life impact resulting from changes in the BCVA of the worse-seeing eye, and presented 6 scenario analyses to take this into account in conjunction with the 3 different sources of utility values:

- Scenario analysis 1: A flat health-related quality of life function where changes in the BCVA of the worse-seeing eye have no impact.
- Scenario analysis 2: A health-related quality of life function where changes in the BCVA of the worse-seeing eye have 15% of the range of changes in the BCVA of the better-seeing eye: that is, a range of 15% of 0.350 which equals 0.053.
- Scenario analysis 3: A health-related quality of life function where changes in the BCVA of the worse-seeing eye have 30% of the range of changes in the BCVA of the better-seeing eye: that is, a range of 30% of 0.350 which equals 0.105.
- Scenario analysis 4: A health-related quality of life function where changes in the BCVA of the worse-seeing eye have 50% of the range of changes in the BCVA of the better-seeing eye: that is, a range of 50% of 0.350 which equals 0.175.
- Scenario analysis 5: A health-related quality of life function where changes in the BCVA of the worse-seeing eye have 70% of the range of changes in the BCVA of the better-seeing eye: that is, a range of 70% of 0.350 which equals 0.245.
- Scenario analysis 6: A health-related quality of life function where changes in the BCVA of the worse-seeing eye have 100% of the range of changes in the BCVA of the better-seeing eye: that is, a range of 0.350.

- 3.39 The ERG also assumed that 20% and 40% of people received unilateral treatment in their better-seeing eye and worse-seeing eye respectively. For the fluocinolone group it was further assumed that 34% of people received bilateral treatment with fluocinolone acetonide intravitreal implant (that is, in both eyes), whereas 6% received treatment with fluocinolone acetonide intravitreal implant in the first eye but not in the second eye. In the optimised standard of care group, the ERG similarly assumed that 20% and 40% of people received unilateral treatment in their better-seeing and worse-seeing eye respectively; whereas 40% of people were assumed to receive treatment in both eyes.
- 3.40 The ERG completed sensitivity analyses without the patient access scheme. Using Brown et al. (1999), the ICERs (based on a 10-letter response criterion and including a bilateral benefit) ranged from £48,533 per QALY gained (scenario analysis 6, where changes in the BCVA of the worse-seeing eye have 100% of the range of changes in the BCVA of the better-seeing eye, that is, a range of 0.350) to £110,730 per QALY gained (scenario analysis 1, where changes in the BCVA of the worse-seeing eye are assumed to have no impact). Using Brown et al. (2000), the equivalent ICERs ranged between £30,910 and £61,942 per QALY gained for fluocinolone acetonide intravitreal implant compared with optimised standard of care. Using the utility values derived from the manufacturer's submission for [NICE technology appraisal guidance 237](#), based upon the 10-letter response criterion and including a bilateral benefit, the ICERs ranged from £69,802 (scenario analysis 6) to £251,686 per QALY gained (scenario analysis 1). Using a 10-letter response criterion, including a bilateral benefit and assuming scenario analysis 3 the ICER values were £47,604 per QALY gained using Brown et al. (2000) utilities and £80,037 per QALY gained using Brown et al. (1999) utilities.
- 3.41 Full details of all the evidence are in the manufacturer's submission and the ERG report for [NICE technology appraisal guidance 271](#).

Rapid review of NICE technology appraisal guidance 271: patient access scheme

- 3.42 In [NICE technology appraisal guidance 271](#), fluocinolone acetonide intravitreal implant was not recommended for treating chronic diabetic macular oedema.

After publication, the manufacturer agreed a patient access scheme with the Department of Health and submitted revised analyses to be considered in a rapid review of the original guidance.

3.43 The manufacturer presented analyses for the full population of people with chronic diabetic macular oedema and for the subgroup of people with chronic diabetic macular oedema who had treatment in an eye with a pseudophakic lens. In the revised economic model the manufacturer included:

- BCVA patient distributions in the extrapolation from month 36 onwards that were specific to a patient's response status
- BCVA patient distributions in the extrapolation for the pseudophakic subgroup that were based on the final 30 months of observations from the FAME trials, and BCVA patient distributions for the full chronic diabetic macular oedema population that were based on the final 12 months of observations
- analyses using utility values from Brown et al. (1999) and Brown et al. (2000), as well as an analysis using utilities from Czoski-Murray et al. (2009) (used in [NICE technology appraisal guidance 237](#), replaced by [technology appraisal guidance 274](#))
- an assumption that changes in vision for people treated in their worse-seeing eye had 30% of the health-related quality of life impact of the same change in vision from treating their best-seeing eye
- an assumption that 20% of patients are unilaterally treated in the best-seeing eye, 40% of patients are unilaterally treated in the worst-seeing eye, and the remaining 40% of patients receive bilateral treatment.

3.44 The manufacturer also clarified some characteristics of the people enrolled in the FAME trials. It stated that people in the trials were as severely affected as patients who would receive fluocinolone acetonide intravitreal implant in routine clinical practice. It stated that in clinical practice, a BCVA of 20/80 is generally considered poor and insufficiently responsive to treatment. It noted that in the FAME trials 59% of the patients with chronic diabetic macular oedema had vision of 20/80 or worse at baseline. Of these, 42.7% and 12.7% of patients had a 15-letter gain in the fluocinolone acetonide intravitreal implant

and optimised standard of care groups respectively ($p < 0.001$). The manufacturer also stated that in the FAME trials all patients had at least 1 prior macular laser treatment before randomisation. It noted that in the optimised standard of care group, people with diabetic macular oedema for more than 3 years showed lower levels of response (13.4% of patients) than people with diabetic macular oedema for less than 3 years (27.8% of patients); whereas in the fluocinolone acetonide intravitreal implant group, people with diabetic macular oedema for more than 3 years showed higher levels of response (34% of patients) than people with diabetic macular oedema for less than 3 years (22.3% of patients). The manufacturer considered that fluocinolone acetonide intravitreal implant provided significant additive benefit for people with chronic diabetic macular oedema whose disease was responding insufficiently to other therapies.

- 3.45 The manufacturer presented the results for the comparison between fluocinolone acetonide intravitreal implant and optimised standard of care with the patient access scheme for the whole population with chronic diabetic macular oedema and for the subgroup of people with a pseudophakic lens. The ICERs for the whole population with the patient access scheme were £37,630 using the utilities from Brown et al. (2000) and £63,472 using the utilities from Brown et al. (1999) per QALY gained. The ICER using the utilities from Czoski-Murray et al. (2009) was £42,663 per QALY gained. In the pseudophakic subgroup the ICERs were £17,639 using the utilities from Brown et al. (2000) and £30,296 using the utilities from Brown et al. (1999) per QALY gained. Using the utilities from Czoski-Murray et al. (2009) the ICER was £19,884 per QALY gained.
- 3.46 The manufacturer carried out a sensitivity analysis assuming that patients would not be re-treated with fluocinolone acetonide intravitreal implant automatically at month 36. Re-treatment would only take place if there was a response to treatment and a BCVA less than 20/32 at month 36. These analyses resulted in small reductions in the ICERs. With the patient access scheme, the ICERs for the chronic diabetic macular oedema population were £34,668 using utilities from Brown et al. (2000) and £57,476 using utilities from Brown et al. (1999) per QALY gained. In the pseudophakic subgroup, the

ICERs were £16,642 using the utilities from Brown et al. (2000) and £28,584 using the utilities from Brown et al. (1999) per QALY gained.

Comments on the manufacturer's rapid review submission by the Evidence Review Group

- 3.47 The ERG stated that there continued to be uncertainty about the best source of utilities to be used in the model. It noted that utilities from Brown et al. (1999) represented a more diverse group of patients with different eye conditions, all of whom were visually impaired with vision of at best 20/40 in at least 1 eye. The ERG considered that utilities from Brown et al. (1999) may be preferable to those from Brown et al. (2000) because Brown et al. (1999) included more patients, which allows a finer gradation of utility estimates for a given BCVA. The number of patients for a given BCVA band was also higher in Brown et al. (1999). The ERG also noted that one-third of patients in Brown et al. (1999) had diabetes compared with none in Brown et al. (2000).
- 3.48 The ERG commented that it was unable to source the utility values from Czoski-Murray et al. (2009) used by the manufacturer. The ERG applied the Czoski-Murray utility function adjusted for an average age of 63 and found similar ranges to those used by the manufacturer. The ERG noted that patients included in Czoski-Murray et al. (2009) used contact lenses to simulate different degrees of visual loss. It also noted that the duration of the simulated visual impairment was short, and so the utility values may not apply to patients with longer duration of visual loss.
- 3.49 The ERG commented that the differences between the ICERs for the whole chronic diabetic macular oedema population and the pseudophakic subgroup were in part driven by differences in the distribution of patients across health states. The ERG noted that there were uncertainties around these distributions because for the pseudophakic subgroup, the difference in the baseline distributions between the fluocinolone group and the optimised standard of care group may indicate a breakdown in randomisation. It also noted that there were differences in the 36-month patient distributions within the optimised standard of care group, between the full chronic diabetic macular oedema population and the pseudophakic subgroup. The ERG commented on the

importance of these patient distributions because they are the basis for the extrapolations in the model.

- 3.50 The ERG checked the manufacturer's ICERs and found minor differences in the values. The ERG commented that there were some small changes made to response rates and drop-out rates that had not been previously applied in the model. The ERG also commented that there were some errors in the adverse effect cost calculations, but correcting for these had little impact on the ICERs. The deterministic ICERs from the ERG check with the patient access scheme for the chronic diabetic macular oedema population were £64,549 and £37,996 per QALY gained using Brown et al. (1999) and Brown et al. (2000) utilities respectively. For the pseudophakic subgroup, the ICERs with the patient access scheme were £30,025 (using Brown et al. [1999] utilities), £21,027 (using Czoski-Murray utilities) and £17,487 (using Brown et al. [2000] utilities) per QALY gained.
- 3.51 The ERG noted that the manufacturer presented a sensitivity analysis that assumed that patients would be re-treated at month 36 only if their disease responded to fluocinolone and they had a BCVA <20/32. The ERG commented that this change only affected costs in the fluocinolone acetonide intravitreal implant group without any clinical impact of these patients stopping treatment. It also stated that the impact on costs may be underestimated as the re-treatment adjustment was only applied to the second re-treatment with fluocinolone acetonide intravitreal implant (at month 36) but not to subsequent re-treatments.

4 Consideration of the evidence

- 4.1 The Appraisal Committee reviewed the data available on the clinical and cost effectiveness of fluocinolone acetonide intravitreal implant, having considered evidence on the nature of chronic diabetic macular oedema after an inadequate response to prior therapy and the value placed on the benefits of fluocinolone acetonide intravitreal implant by people with the condition, those who represent them, and clinical specialists. It also took into account the effective use of NHS resources.
- 4.2 The Committee discussed the causes of diabetic macular oedema, and recognised the importance of good control of blood glucose, blood pressure and blood lipids in reducing the risk of diabetic macular oedema, the progression of diabetic macular oedema and other complications. The Committee discussed the impact of visual impairment on people with diabetic macular oedema. It was aware of comments from patient experts describing the significant negative impact that visual impairment has on the physical and emotional wellbeing of people with diabetic macular oedema. The Committee noted that people with diabetes manage some aspects of their own condition and that visual impairment can affect their ability to do this. This in turn can increase the risk of long-term disease complications such as kidney disease, cardiovascular disease and amputations. The Committee also heard from clinical specialists that chronic diabetic macular oedema tends to be a bilateral condition (that is, affecting both eyes) with the potential risk of losing sight in both eyes. The clinical specialists stated that chronic diabetic macular oedema therefore affects the quality of life of people with the condition by causing visual impairment and also that of their carers because of loss of independence. The Committee noted comments from clinical specialists which highlighted that several studies have reported depression in people with chronic diabetic macular oedema. The Committee recognised the impact of visual impairment on people with diabetic macular oedema, particularly those with chronic diabetic macular oedema that is not responsive to available therapies.
- 4.3 The Committee discussed the current management of diabetic macular oedema. The Committee was aware that [Ranibizumab for treating diabetic](#)

macular oedema (NICE technology appraisal guidance 274) is the only NICE guidance relating specifically to treating diabetic macular oedema. It noted that the marketing authorisation for fluocinolone acetonide intravitreal implant specified its use only when diabetic macular oedema was insufficiently responsive to available therapies and that this population was not covered by NICE technology appraisal guidance 274. The Committee heard from the clinical specialists that when anti-vascular endothelial growth factor (anti-VEGF) therapies are not options, standard treatment would be 'optimised standard of care', which could include laser therapy as a maintenance treatment. The Committee was aware of the publication of guidelines from the Royal College of Ophthalmologists suggesting the use of fluocinolone for some people and understood that there are currently no standard treatments for people with chronic diabetic macular oedema after other therapies have failed.

- 4.4 The Committee discussed the likely place of fluocinolone acetonide intravitreal implant in clinical practice. It heard from the clinical specialists that because fluocinolone acetonide intravitreal implant is a sustained-release low-dose long-acting steroid, it has clear advantages over other steroid implants, which are not licensed for the treatment of diabetic macular oedema. However, the Committee understood that there are significant side effects associated with the use of a steroid in the eye, especially the acceleration of cataract development and increased rates of raised intraocular pressure, and that these still occur with fluocinolone acetonide intravitreal implant. The Committee discussed whether the use of fluocinolone acetonide intravitreal implant in one eye could be associated with cataract development in the other eye. It understood from the manufacturer that it was not aware of any evidence suggesting this. Further, it heard that fluocinolone was not detectable in the blood. The Committee also heard from the clinical specialists that there is a spectrum of severity of diabetic macular oedema, and steroid treatment might be useful in those with more severe diabetic macular oedema if other treatments (including laser therapy) have failed and cataracts have already been removed. It understood that based on existing NICE guidance and clinical practice the patient population that would be considered for fluocinolone acetonide intravitreal implant would be those for whom laser photocoagulation and anti-VEGF therapies had failed.

Clinical effectiveness

- 4.5 The Committee considered the evidence presented by the manufacturer on the clinical effectiveness of fluocinolone acetonide intravitreal implant. The Committee noted that the main sources of evidence were the FAME A and B randomised controlled trials, which enrolled people with diabetic macular oedema and included a preplanned analysis defined by the median duration of diabetic macular oedema, which had been used to identify a subgroup of people with chronic diabetic macular oedema (that is, of more than 3 years' duration). It noted that in the FAME trials, patients were only treated in 1 eye and that the majority of patients were treated in their worse-seeing eye. The Committee noted that fluocinolone acetonide intravitreal implant was associated with statistically significant gains in the proportion of patients with chronic diabetic macular oedema who had a more than 15-letter increase in best corrected visual acuity (BCVA) from baseline compared with the sham injection group. It also noted that mean change in BCVA from baseline at month 36 of the FAME trials was statistically significantly greater in the fluocinolone acetonide intravitreal implant group compared with the sham injection group. The Committee concluded that fluocinolone acetonide intravitreal implant showed greater clinical effectiveness than sham injection in people with chronic diabetic macular oedema.
- 4.6 The Committee discussed the characteristics of the FAME trial population and how these related to the patient group specified in the marketing authorisation for fluocinolone acetonide intravitreal implant and to the group of patients likely to receive treatment in UK clinical practice. The Committee noted that the efficacy data submitted were based on a subgroup defined as having chronic disease with a duration of longer than 3 years, rather than as unresponsive to available therapies as specified in the marketing authorisation. The Committee further noted that some patients included received additional treatments during the trials such as laser photocoagulation and anti-VEGF injections. The Committee discussed the concerns that this implied that in the FAME trials fluocinolone acetonide intravitreal implant was not being used as specified in the marketing authorisation. The Committee noted the manufacturer's submission and comments on the appraisal consultation document stating that the patients in the FAME trials had poor visual acuity, and that the patients in

the comparator arm had disease that was responding poorly to the background therapies when given in the trial. However, the Committee remained concerned that the data from the FAME trials did not specifically reflect a group of patients who had disease that was insufficiently responsive to other therapies because people in the FAME trials did not necessarily have disease that was unresponsive to treatment with anti-VEGF or laser photocoagulation before randomisation. The Committee considered that this could affect the levels of response observed in clinical practice because one of the mechanisms of action of corticosteroids was to act on VEGF. It noted comments received on the appraisal consultation document that studies of other intravitreal corticosteroids (such as triamcinolone) have shown a beneficial effect on macular oedema when used after anti-VEGF treatments. The Committee remained concerned that the data from the trials may not be fully representative of the group of people who would be receiving fluocinolone acetonide intravitreal implant in UK clinical practice. However, it concluded that although there is uncertainty about the generalisability of the trial data to UK clinical practice, fluocinolone acetonide intravitreal implant is likely to be clinically effective in this population.

- 4.7 The Committee considered the manufacturer's approach to conducting an indirect comparison of fluocinolone acetonide intravitreal implant and laser photocoagulation using the FAME trials and the DRCR Protocol B study (2008). The Committee was aware of the Evidence Review Group's (ERG's) concerns over the value of the indirect comparison given the absence of a common comparator to link the FAME studies with the identified DRCR Protocol B study. It also noted that the retinopathy in patients included in the DRCR study was not as severe as in the FAME trials; that 40% of the DRCR population was laser-naïve; and that in the DRCR study there was no stipulation on duration of diabetic macular oedema at randomisation. The Committee concluded that the indirect comparison could not be interpreted with confidence and in any case was inappropriate to the scope for this appraisal, which considered fluocinolone acetonide intravitreal implant when diabetic macular oedema has been insufficiently responsive to available therapies.

- 4.8 The Committee considered the evidence for adverse events associated with fluocinolone acetonide intravitreal implant. It was aware that fluocinolone acetonide intravitreal implant was associated with the formation or progression of cataract. It noted that although there were similar rates of pre-existing cataract between the 2 trial groups (77.9% and 77.0%) among people who were phakic (still had their natural lens) at baseline, cataract surgery was needed by a greater percentage (85.1% versus 36.4%) of these people in the fluocinolone acetonide intravitreal implant group than in the sham injection group. The Committee heard from clinical specialists that the majority of people with chronic diabetic macular oedema would be likely to develop cataracts at some stage but that fluocinolone acetonide intravitreal implant might accelerate this. The Committee also noted that administration of fluocinolone acetonide intravitreal implant was associated with increased intraocular pressure: 5.3% of people in the fluocinolone group needed intraocular pressure-lowering surgery compared with 0% in the sham injection group. The Committee heard from clinical specialists that intraocular pressure-lowering surgery was a particular concern and for this reason, in addition to the associated acceleration of cataract development, clinicians would be likely to use fluocinolone acetonide intravitreal implant conservatively and be reluctant to use it too early in the treatment pathway. The Committee concluded that it was appropriate to take account of these adverse events when considering the approach to economic modelling.

Cost effectiveness

- 4.9 The Committee considered the manufacturer's economic model and sensitivity analyses, and discussed the key parameters used in it. The Committee noted the sensitivity of the model to the assumptions about the relationship between a person's treated eye and their overall visual acuity, the number of fluocinolone acetonide intravitreal implant treatments in the first 3 years, assumptions about the benefits of treatment after month 36, the rate of re-treatment, and the source of health-related quality of life values. The Committee concluded that the cost-effectiveness estimates were most sensitive to the source of health-related quality of life values, the assumption that a person's overall visual acuity related only to their treated eye, and the assumption that all treated eyes were better-seeing eyes.

- 4.10 The Committee considered the ERG's critique of the manufacturer's original base-case results. It was aware that the ERG had made a series of explorations using the manufacturer's original base-case incremental cost-effectiveness ratio (ICER) of £22,600 per quality-adjusted life year (QALY) gained as a starting point. The Committee noted that by correcting an error relating to the application of annual discontinuation rates in the model, and by changing the formulae for averaging male and female mortality rates and for calculating the percentage of patients remaining on fluocinolone acetonide intravitreal implant beyond 9 years, the ERG's revisions increased the manufacturer's original base-case ICER from £22,700 to £26,500 per QALY gained for fluocinolone acetonide intravitreal implant compared with optimised standard of care without the patient access scheme. The Committee agreed that the ERG's initial error corrections to the model were appropriate.
- 4.11 The Committee went on to consider the ERG's further explorations to the original model, which were included in the manufacturer's revised model submitted for the rapid review:
- aligning assumptions about the rate of laser administrations per patient in the optimised standard of care group for year 1 of the model with the trial data
 - applying revised unit costs for some adverse events
 - revising the assumptions about the extrapolation of benefits beyond 3 years in the model
 - applying an adjusted bilateral treatment rate in the fluocinolone acetonide intravitreal implant group of 85.2% of the treatment rate in the optimised standard of care group
 - use of a 10-letters response criterion
 - applying the absolute cost and quality of life uplift associated with bilateral treatment in the optimised standard of care group to the patients in the fluocinolone group needing bilateral treatment but for whom it was contraindicated
 - applying the cost of blindness when only the treated eye fell below a BCVA of 35 letters rather than both eyes.

The Committee concluded that the above amendments to the assumptions in the economic model submitted by the manufacturer for the rapid review were reasonable.

- 4.12 The Committee discussed the assumption in the model that 20% of patients were treated in their best-seeing eye, 40% in their worse-seeing eye and 40% of patients were treated in both eyes. The Committee noted the manufacturer's comment on the appraisal consultation document that in the FAME trials the majority of the patients had their worse-seeing eye treated because of ethical considerations but that in clinical practice, more patients would be treated in their best-seeing eye because this is more likely to protect functional vision. The Committee agreed that this could be the case but considered that currently there is a lack of evidence. It concluded that the assumption of 20% of the patients being treated in their best-seeing eye, 40% in their worse-seeing eye and 40% of patients receiving bilateral treatment was reasonable.
- 4.13 The Committee discussed the use of second and subsequent fluocinolone acetonide intravitreal implants and the assumption in the model of re-treatment only after 36 months. It noted that the summary of product characteristics states that an additional implant may be administered after 12 months if the patient experiences decreased vision or an increase in retinal thickness secondary to recurrent or worsening diabetic macular oedema. The Committee heard from the manufacturer that this wording was included in the summary of product characteristics because patients in the FAME trials could receive a further implant after 12 months. However, subsequent research showed that the effect of the corticosteroid should last 36 months. The Committee understood from the manufacturer that it does not promote having the implants more frequently than once every 36 months and that there are currently no data available demonstrating further benefit from reimplantation. The Committee accepted the assumption in the model of no more than 1 implant every 36 months.
- 4.14 The Committee discussed the fact that the model did not apply quality of life detriments to adverse events such as cataracts, glaucoma or raised intraocular pressure. The Committee noted the manufacturer's rationale that the impact of

adverse events such as cataract formation on visual acuity was incorporated in the overall utility measures, that the application of these decrements would be short in duration and therefore have a minimal impact on the ICERs, and that the decrements associated with surgery would depend on when a patient chose to have surgery. Nevertheless, the Committee concluded that if disutilities associated with operations, procedures and hospital attendances (such as cataract removal, glaucoma surgery, retinal detachment surgery, vitrectomy and treatment of endophthalmitis) had been taken into account the ICERs would increase.

- 4.15 The Committee discussed the utility values for the better-seeing eye used in the economic model. The Committee noted that the manufacturer's original analyses had been based on Brown et al. (2000) utilities. It further noted that the ERG's explorations of the manufacturer's original model retained the Brown et al. (2000) utility values, and also explored the effect of applying utility values from Brown et al. (1999) and from a regression analysis in the manufacturer's submission for [NICE technology appraisal guidance 237](#). The Committee agreed that the utility values may vary in their appropriateness in being applied to people with chronic diabetic macular oedema. It accepted that the utility values from the manufacturer's submission for NICE technology appraisal guidance 237 were relatively insensitive to changes in visual acuity and therefore ICERs using these utilities may be numerically too high.
- 4.16 The Committee discussed further the relative merits of Brown et al. (1999) and Brown et al. (2000). It was aware that the Brown et al. (1999) utility values were based on a larger population with a mix of eye diseases whereas the Brown et al. (2000) study was based on a smaller population with age-related macular oedema (but not diabetic macular oedema). It noted that the data from Brown et al. (2000) included a large drop in utility values between some health states, which may have arisen from the smaller patient numbers in each health state. It also noted that utility values from Brown et al. (2000) for the best health states were higher than would be expected for people with chronic diabetic macular oedema, who are usually older people with comorbidities and this, in turn, increased the uncertainty around the validity of this source of utility values. The Committee heard from the ERG that Brown et al. (2002) found no difference in the reduction in the quality of life associated with similar levels of

visual acuity loss depending on the causes of vision impairment. The Committee considered the manufacturer's rationale for using Brown et al. (2000) values based on the higher proportion of patients with oedema compared with Brown et al. (1999). It also noted that the manufacturer restated in its comment on the appraisal consultation document that Brown et al. (2000) represents a more accurate source of utility values for people with diabetic macular oedema. However, the Committee was not persuaded that this outweighed the benefits of the larger sample size and the associated finer delineation of utilities possible with larger patient numbers. The Committee concluded that there were limitations to using the available utility values to model the health-related quality of life of the group of people with chronic diabetic macular oedema. However, of the available source of utility values, Brown et al. (1999) had advantages because of the size of the sample included in the study. Because there were some uncertainties about the most appropriate source of utility values for this patient population the Committee agreed to consider a range of ICERs based on both Brown et al. (1999) and Brown et al. (2000).

- 4.17 The Committee noted that the manufacturer had also included analyses using utility values from Czoski-Murray et al. (2009) and Heintz et al. (2012). The Committee discussed the appropriateness of using the values from Heintz et al. (2012) and considered that the very slight differences between utilities for the loss of better-seeing eye vision relative to worse-seeing eye vision lacked face validity. The Committee then discussed the utility values presented in Czoski-Murray (2009). It understood that the source of utility values had been considered for the appraisal of [Ranibizumab for the treatment of diabetic macular oedema](#) (NICE technology appraisal guidance 237, replaced by [NICE technology appraisal guidance 274](#)) and noted that using these utility values resulted in ICERs in between the ones obtained using Brown et al. (1999) and Brown et al. (2000). The Committee agreed to consider Czoski-Murray et al. (2009) as another source of utility values for this appraisal.
- 4.18 The Committee discussed the most appropriate adjustments needed to the better-seeing eye utility values in the model when the worse-seeing eye or both eyes were treated. The Committee considered the 6 scenario analyses carried out by the ERG (see [section 3.38](#)) which varied the health-related

quality of life impact of changes in the vision of the worse-seeing eye and the resultant QALY gain associated with treatment of the worse-seeing eye or both eyes. The Committee understood the concerns of the ERG about using the Heintz et al. (2012) data to calculate the amount of gain from treating the worse-seeing eye, because of the small numbers of patients in some of the visual acuity levels and also because the manufacturer's calculation did not account for visual acuity in the better-seeing eye being correlated with visual acuity in the worse-seeing eye. The Committee also discussed the utilities for worse-seeing eyes in Brown et al. (1999) and heard from the ERG about an additional study, Sahel et al. (2007), which also showed that the visual acuity in the worse-seeing eye has little impact on health-related quality of life unless it is severely affected. The Committee was aware that there could be psychological benefits from treating the worse-seeing eye that had not been captured in the calculation of the QALY. However, the Committee noted that the FAME trials collected data (with a majority of worse-seeing eyes) on the effect of visual impairment on quality of life using the disease-specific VFQ-25 questionnaire, and that this had shown no difference between the groups treated with fluocinolone acetonide intravitreal implant and those treated with sham injection. On balance, the Committee considered scenario analysis 3 (which assumed that changes in vision for people treated in their worse-seeing eye had 30% of the health-related quality of life impact of the same change in vision from treating their better-seeing eye) to be the most appropriate for decision-making. The Committee was not persuaded that 30% was an unreasonable reflection of the clinical situation for people with chronic diabetic macular oedema and noted that this was also consistent with previous appraisals.

- 4.19 The Committee then considered the cost-effectiveness results based on data from the FAME trials for all patients with chronic diabetic macular oedema taking into account its concerns (sections 4.9–4.18). The Committee noted that the most plausible ICER for fluocinolone acetonide intravitreal implant compared with optimised standard of care in the original guidance was at least £47,600 using the utilities from Brown et al. (2000) and £80,000 using the utilities from Brown et al. (1999) per QALY gained. The Committee noted that with the patient access scheme the ICERs were reduced to £37,600 using the utilities from Brown et al. (2000), £42,700 using the utilities from Czoski-Murray

et al. (2009) and £63,500 using the utilities from Brown et al. (1999) per QALY gained. The Committee considered that the lowest of these estimates remained over £30,000 per QALY gained and therefore outside the range normally considered cost effective (£20,000–£30,000 per QALY gained). It also noted that there was substantial uncertainty in these estimates, particularly in the extent to which the results of the clinical trial data included in the analyses could be applied to the population of people who would receive treatment in clinical practice. The Committee concluded that fluocinolone acetonide intravitreal implant could not be recommended as a cost-effective use of NHS resources for treating people with chronic diabetic macular oedema that is insufficiently responsive to available therapies.

- 4.20 The Committee discussed the subgroup of people with chronic diabetic macular oedema who had treatment in an eye with a pseudophakic lens. The Committee acknowledged that this is an identifiable subgroup of people with chronic diabetic macular oedema. The Committee noted that the numbers of patients in the subgroup were approximately half the chronic diabetic macular oedema population in the FAME trials and that the clinical effectiveness in terms of 15-letter gain in visual acuity in this subgroup was numerically worse than that in the total population (albeit with wide confidence intervals). The Committee considered that the comparatively small numbers of such patients in the FAME trials led to uncertainty in the estimates of clinical effectiveness for this group and thus in the estimates from the economic modelling but accepted that the subgroup proposed was reasonable. The Committee also considered that its concerns about the extent to which the data for the subgroup were representative of the group of people who would receive fluocinolone acetonide intravitreal implants in UK clinical practice remained valid (section 4.6). However, on balance, the Committee concluded that fluocinolone acetonide intravitreal implant had been shown to be clinically effective in this subgroup of people.
- 4.21 The Committee considered the cost effectiveness of fluocinolone acetonide intravitreal implant in the pseudophakic subgroup. It noted that the most plausible ICER for this subgroup in [NICE technology appraisal guidance 271](#) was between £29,700 per QALY gained using the utilities from Brown et al. (2000) and £50,600 using the utilities from Brown et al. (1999). The Committee

noted that with the patient access scheme, the ICERs presented by the ERG were £30,000 per QALY gained using the utilities from Brown et al. (1999), £21,000 per QALY gained using the utilities from Czoski-Murray et al. (2009) and £17,500 per QALY gained using the utilities from Brown et al. (2000). The Committee considered that the most plausible estimates of cost effectiveness would be in the upper end of this range, and that there was significant uncertainty around this estimate. The Committee was persuaded that the technology had been shown to meet a clinical need in people whose disease is unresponsive to available therapies. On balance, the Committee concluded that fluocinolone acetonide intravitreal implant could be a cost-effective use of NHS resources and recommended it as an option for people with chronic diabetic macular oedema that is insufficiently responsive to available therapies and if the implant is to be used in an eye with an intraocular (pseudophakic) lens.

Summary of Appraisal Committee's key conclusions

TA301	Appraisal title: Fluocinolone acetonide intravitreal implant for treating chronic diabetic macular oedema after an inadequate response to prior therapy (rapid review of technology appraisal 271)	Section
Key conclusion		

Fluocinolone acetonide intravitreal implant is recommended as an option for treating chronic diabetic macular oedema that is insufficiently responsive to available therapies only if the implant is to be used in an eye with an intraocular (pseudophakic) lens and the manufacturer provides fluocinolone acetonide intravitreal implant with the discount agreed in the patient access scheme. The Committee considered that the lowest of ICERs with the patient access scheme for people with chronic diabetic macular oedema remained over £30,000 per QALY gained and therefore outside the range normally considered cost-effective (£20,000–£30,000 per QALY gained). The Committee concluded that fluocinolone acetonide intravitreal implant was not recommended as a cost-effective use of NHS resources for treating people with chronic diabetic macular oedema that is insufficiently responsive to available therapies. The Committee noted that for the pseudophakic subgroup, the ICERs with the patient access scheme checked by the ERG were between £30,000 per QALY gained using the utilities from Brown et al. (1999) and £17,500 per QALY gained with the utilities from Brown et al. (2000). It was persuaded that the technology had been shown to meet a clinical need in people whose disease is unresponsive to available therapies. The Committee concluded that fluocinolone acetonide intravitreal implant was a cost-effective use of NHS resources.		1.1, 4.19, 4.21
Current practice		
Clinical need of patients, including the availability of alternative treatments	The Committee was aware of comments from patient experts describing the significant negative impact that visual impairment has on the physical and emotional wellbeing of people with diabetic macular oedema. The Committee was aware that Ranibizumab for treating diabetic macular oedema (NICE technology appraisal guidance 274) is the only NICE guidance relating specifically to treating diabetic macular oedema and that it does not cover the treatment of people with unresponsive disease. The Committee understood that there are currently no standard treatments for people with chronic diabetic macular oedema after other therapies have failed.	4.2, 4.3
The technology		

Proposed benefits of the technology How innovative is the technology in its potential to make a significant and substantial impact on health-related benefits?	The Committee heard from the clinical specialists that because fluocinolone acetonide intravitreal implant is a sustained-release low-dose long-acting steroid, it has clear advantages over other steroid implants, which the Committee was aware are not licensed for the treatment of diabetic macular oedema.	4.4
What is the position of the treatment in the pathway of care for the condition?	The Committee noted that the marketing authorisation for fluocinolone acetonide intravitreal implant specified its use only when diabetic macular oedema was insufficiently responsive to available therapies and that this population was not covered by NICE technology appraisal guidance 274 . It understood that based on existing NICE guidance and clinical practice the patient population that would be considered for fluocinolone acetonide intravitreal implant would be those for whom laser photocoagulation and anti-VEGF therapies had failed.	4.3, 4.4
Adverse reactions	The Committee noted that the significant side effects associated with the use of a steroid in the eye, especially the acceleration of cataract development and increased rates of raised intraocular pressure, still occur with fluocinolone acetonide intravitreal implant.	4.8
Evidence for clinical effectiveness		
Availability, nature and quality of evidence	The Committee noted that the main sources of evidence were the FAME A and B randomised controlled trials. The Committee noted the DRCR Protocol B study used by the manufacturer in its indirect comparison. The Committee was aware of the ERG's concerns about the indirect comparison and concluded that it could not be interpreted with confidence and was inappropriate to the scope of this appraisal.	4.5, 4.7

Relevance to general clinical practice in the NHS	The Committee noted that the efficacy data submitted were based on a subgroup defined as having chronic disease with a duration of longer than 3 years, rather than as unresponsive to available therapies as specified in the marketing authorisation. The Committee further noted that some patients included in the trials received additional treatments during the trial such as laser photocoagulation and anti-VEGF injections, and that this implied that in the FAME trials fluocinolone acetonide intravitreal implant was not being used as specified in the marketing authorisation. However, it concluded that although there is uncertainty about the generalisability of the trial data to UK clinical practice, fluocinolone acetonide intravitreal implant is likely to be clinically effective in this population.	4.6
Uncertainties generated by the evidence	The Committee was concerned that the data from the FAME trials did not reflect a group of patients who had disease that was insufficiently responsive to other therapies because people in the FAME trials did not necessarily have disease that was unresponsive to treatment with anti-VEGF or laser photocoagulation before randomisation.	4.6
Are there any clinically relevant subgroups for which there is evidence of differential effectiveness?	The Committee discussed the subgroup of people with chronic diabetic macular oedema who had treatment in an eye with a pseudophakic lens. The Committee considered that the comparatively small numbers of such patients in the FAME trials led to uncertainty in the estimates of clinical effectiveness for this group and thus in the estimates from the economic modelling but accepted that the subgroup proposed was reasonable.	4.20
Estimate of the size of the clinical effectiveness including strength of supporting evidence	The Committee concluded that fluocinolone acetonide intravitreal implant showed greater clinical effectiveness than sham injection in people with chronic diabetic macular oedema. The Committee concluded that fluocinolone acetonide intravitreal implant had been shown to be clinically effective in the subgroup of people who had a pseudophakic lens.	4.5, 4.20

Evidence for cost effectiveness		
Availability and nature of evidence	The Committee considered the manufacturer's model and the ERG's critique submitted for the rapid review.	4.9, 4.10
Uncertainties around and plausibility of assumptions and inputs in the economic model	The Committee concluded that the manufacturer's model did not take into account the disutilities associated with operations, procedures and hospital attendances which if taken into account would cause the ICER to be even higher. The Committee concluded that there were limitations to using the available utilities to model the health-related quality of life of the group of people with chronic diabetic macular oedema. The Committee considered that the comparatively small numbers of patients who had a pseudophakic lens in the FAME trials led to uncertainty in the estimates of clinical effectiveness for this group and thus in the estimates from the economic modelling.	4.14, 4.16, 4.20
Incorporation of health-related quality-of-life benefits and utility values Have any potential significant and substantial health-related benefits been identified that were not included in the economic model, and how have they been considered?	The Committee concluded that there were limitations to using the available utilities to model the health-related quality of life of the group of people with chronic diabetic macular oedema and agreed to consider a range of ICERs based on both Brown et al. (1999) and Brown et al. (2000). The Committee noted that the manufacturer had also included analyses using utility values from Czoski-Murray et al. (2009) and Heintz et al. (2012). The Committee discussed the appropriateness of using the values from Heintz et al. (2012) and considered that the very slight differences between utilities for the loss of better-seeing eye vision relative to worse-seeing eye vision lacked face validity. It understood that the utility values from Czoski-Murray et al. (2009) had been considered for the appraisal of ranibizumab for the treatment of diabetic macular oedema and agreed to consider it as another source of utility values for this appraisal.	4.16, 4.17

Are there specific groups of people for whom the technology is particularly cost effective?	The Committee discussed the subgroup of people with chronic diabetic macular oedema who had treatment in an eye with a pseudophakic lens and concluded that it was a cost-effective use of NHS resources.	4.21
What are the key drivers of cost effectiveness?	The Committee concluded that the cost-effectiveness estimates were most sensitive to the source of health-related quality of life values and the assumption that a person's overall visual acuity related only to their treated eye.	4.9
Most likely cost-effectiveness estimate (given as an ICER)	The Committee noted that for all patients with chronic diabetic macular oedema, with the patient access scheme, the ICERs were £37,600 using the utilities from Brown et al. (2000), £42,700 using the utilities from Czoski-Murray et al. (2009) and £63,500 using the utilities from Brown et al. (1999) per QALY gained. The Committee noted that for the pseudophakic subgroup, the ICERs with the patient access scheme presented by the ERG were £30,000 per QALY gained using the utilities from Brown et al. (1999), £21,000 per QALY gained with the utilities from Czoski-Murray et al. (2009) and £17,500 per QALY gained with the utilities from Brown et al. (2000).	4.19, 4.21
Additional factors taken into account		
Patient access schemes (PPRS)	The manufacturer of fluocinolone acetonide intravitreal implant has agreed a patient access scheme with the Department of Health in which fluocinolone acetonide intravitreal implant will be available with a discount (see section 5.3). The details of the scheme were provided as commercial in confidence. The Department of Health considered that this patient access scheme does not constitute an excessive administrative burden on the NHS.	2.4
End-of-life considerations	Not applicable.	

Equalities considerations and social value judgements	The Committee concluded there were no issues relating to the equalities legislation, and there was no need to alter or add to its recommendations.	

5 Implementation

- 5.1 Section 7(6) of the [National Institute for Health and Care Excellence \(Constitution and Functions\)](#) and the [Health and Social Care Information Centre \(Functions\) Regulations 2013](#) requires clinical commissioning groups, NHS England and, with respect to their public health functions, local authorities to comply with the recommendations in this appraisal within 3 months of its date of publication.
- 5.2 When NICE recommends a treatment 'as an option', the NHS must make sure it is available within the period set out in the paragraph above. This means that, if a patient has diabetic macular oedema and has a pseudophakic lens and the doctor responsible for their care thinks that fluocinolone acetonide intravitreal implant is the right treatment, it should be available for use, in line with NICE's recommendations.
- 5.3 The Department of Health and the manufacturer have agreed that fluocinolone acetonide intravitreal implant will be available to the NHS with a patient access scheme which makes fluocinolone acetonide intravitreal implant available with a discount. The size of the discount is commercial in confidence. It is the responsibility of the manufacturer to communicate details of the discount to the relevant NHS organisations. Any enquiries from NHS organisations about the patient access scheme should be directed to Dr Eric Teo, eric.teo@alimerasciences.com, 01252 761203.
- 5.4 NICE has developed a [costing template and report](#) to estimate the national and local savings and costs associated with implementation to help organisations put this guidance into practice.

6 Recommendations for further research

- 6.1 Further research is recommended to resolve uncertainties about the cost effectiveness of fluocinolone acetonide intravitreal implant for the treatment of chronic diabetic macular oedema. This should focus on a group of patients whose condition is unresponsive to other available therapies, and include measures of efficacy and health-related quality of life. Research should focus on identifying appropriate utility values, taking into account the utility values for different levels of visual acuity and the relative relationship in utility values from treating the best-seeing and the worse-seeing eye. The appropriateness of generalising utility values from one group of eye conditions to another group would also be of value.

7 Related NICE guidance

Details are correct at the time of publication. Further information is available on the [NICE website](#).

Published

- [Aflibercept solution for injection for treating wet age-related macular degeneration](#). NICE technology appraisal guidance 294 (2013).
- [Ranibizumab for treating visual impairment caused by macular oedema secondary to retinal vein occlusion](#). NICE technology appraisal guidance 283 (2013).
- [Ranibizumab for treating diabetic macular oedema \(rapid review of technology appraisal guidance 237\)](#) . NICE technology appraisal guidance 274 (2013).
- [Dexamethasone intravitreal implant for the treatment of macular oedema secondary to retinal vein occlusion](#). NICE technology appraisal guidance 229 (2011).
- [Type 2 diabetes: the management of type 2 diabetes](#). NICE clinical guideline 87 (2009).
- [Ranibizumab and pegaptanib for the treatment of age-related macular degeneration](#). NICE technology appraisal guidance 155 (2008).
- [Type 1 diabetes](#). NICE clinical guideline 15 (2004).

8 Review of guidance

- 8.1 The guidance on this technology will be considered for review by the Guidance Executive in November 2016. The Guidance Executive will decide whether the technology should be reviewed based on information gathered by NICE, and in consultation with consultees and commentators.

Andrew Dillon
Chief Executive
November 2013

9 Appraisal Committee members, guideline representatives and NICE project team

9.1 Appraisal Committee members

The Appraisal Committees are standing advisory committees of NICE. Members are appointed for a 3-year term. A list of the Committee members who took part in the discussions for this appraisal appears below. There are 4 Appraisal Committees, each with a chair and vice chair. Each Appraisal Committee meets once a month, except in December when there are no meetings. Each Committee considers its own list of technologies, and ongoing topics are not moved between Committees.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

The minutes of each Appraisal Committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Professor Andrew Stevens

Chair of Appraisal Committee C, Professor of Public Health, University of Birmingham

Professor Eugene Milne

Vice Chair of Appraisal Committee C, Deputy Regional Director of Public Health, North East Strategic Health Authority, Newcastle upon Tyne

Dr David Black

Medical Director, NHS South Yorkshire and Bassetlaw

Dr Daniele Bryden

Consultant in Intensive Care Medicine and Anaesthesia, Sheffield Teaching Hospitals NHS Trust

Dr Andrew Burnet

Formerly - Director for Health Improvement and Medical Director, NHS Barnet, London

David Chandler

Lay Member

Gail Coster

Advanced Practice Sonographer, Mid Yorkshire Hospitals NHS Trust

Professor Peter Crome

Honorary Professor, Dept of Primary Care and Population Health, University College London

Dr Maria Dyban

General Practitioner, Kings Road Surgery, Cardiff

Professor Rachel A Elliott

Lord Trent Professor of Medicines and Health, University of Nottingham

Dr Greg Fell

Consultant in Public Health, Bradford Metropolitan Borough Council

Dr Wasim Hanif

Consultant Physician and Honorary Senior Lecturer, University Hospital Birmingham

Dr Alan Haycox

Reader in Health Economics, University of Liverpool Management School

Dr Peter Jackson

Clinical Pharmacologist, University of Sheffield

Dr Janice Kohler

Senior Lecturer and Consultant in Paediatric Oncology, Southampton University Hospital Trust

Emily Lam

Lay Member

Dr Allyson Lipp

Principal Lecturer, University of South Wales

Dr Claire McKenna

Research Fellow in Health Economics, University of York

Professor Gary McVeigh

Professor of Cardiovascular Medicine, Queens University Belfast and Consultant Physician, Belfast City Hospital

Dr Grant Maclaine

Director, Health Economics & Outcomes Research, BD, Oxford

Dr Andrea Manca

Health Economist and Senior Research Fellow, University of York

Henry Marsh

Consultant Neurosurgeon, St George's Hospital, London

Dr Paul Miller

Director, Payer Evidence, AstraZeneca UK Ltd

Professor Stephen O'Brien

Professor of Haematology, Newcastle University

Alan Rigby

Academic Reader, University of Hull

Dr Peter Selby

Consultant Physician, Central Manchester University Hospitals NHS Foundation Trust

Professor Matt Stevenson

Technical Director, School of Health and Related Research, University of Sheffield

Dr Paul Tappenden

Reader in Health Economic Modelling, School of Health and Related Research, University of Sheffield

Dr Judith Wardle

Lay Member

9.2 NICE project team

Each technology appraisal is assigned to a team consisting of 1 or more health technology analysts (who act as technical leads for the appraisal), a technical adviser and a project manager.

Pilar Pinilla-Dominguez

Technical Lead

Zoe Garrett

Technical Adviser

Lori Farrar

Project Manager

10 Sources of evidence considered by the Committee

A The Evidence Review Group (ERG) report for [Fluocinolone acetonide intravitreal implant for the treatment of chronic diabetic macular oedema after an inadequate response to prior therapy](#) (NICE technology appraisal 271) was prepared by Warwick Evidence:

- Cummins E, Shyangdan D, Royle P, Johnston R, Ho V, Briscoe S, Lois N, Waugh N. Fluocinolone acetonide intravitreal implant for the treatment of chronic diabetic macular oedema: A Single Technology Appraisal. Warwick Evidence, 2012

The critique for this appraisal was prepared by Warwick Evidence:

- Cummins E, Lois N. Waugh N. Fluocinolone for diabetic macular oedema: evidence review group report on Alimera submission for rapid review of TA 271.

B The following organisations accepted the invitation to participate in this appraisal as consultees and commentators. They were invited to comment on the draft scope, the ERG report and the appraisal consultation document (ACD). Organisations listed in I were also invited to make written submissions. Organisations listed in II and III had the opportunity to give their expert views. Organisations listed in I, II and III also have the opportunity to appeal against the final appraisal determination.

I Manufacturer/sponsor:

- Alimera Sciences Inc

II Professional/specialist and patient/carers groups:

- Diabetes UK
- Fight for Sight
- Organisation of Blind African Caribbeans
- Royal National Institute of Blind People (RNIB)
- Royal College of Nursing
- Royal College of Ophthalmologists

- Royal College of Physicians

III Other consultees:

- Department of Health
- Welsh Government

IV Commentator organisations (did not provide written evidence and without the right of appeal):

- Commissioning Support Appraisals Service
- Department of Health, Social Services and Public Safety for Northern Ireland
- Healthcare Improvement Scotland
- Novartis Pharmaceuticals
- Pfizer
- Roche Products

C The following individuals were selected from clinical specialist and patient expert nominations from the non-manufacturer/sponsor consultees and commentators. They gave their expert personal view on Fluocinolone acetonide intravitreal implant for the treatment of chronic diabetic macular oedema by attending the initial Committee discussion and providing written evidence to the Committee. They are invited to comment on the ACD.

- Clare Bailey, Consultant Ophthalmologist, nominated by Alimera Sciences Inc – clinical specialist
- Professor Yit Yang, Consultant Ophthalmologist, Nominated by Royal College of Ophthalmologists – clinical specialist
- Clara Eaglen, Policy and Campaigns manager, nominated by Royal National Institute of Blind People – patient expert

D Representatives from the following manufacturer/sponsor attended Committee meetings. They contributed only when asked by the Committee chair to clarify specific issues and comment on factual accuracy.

- Alimera Sciences Inc

About this guidance

NICE technology appraisal guidance is about the use of new and existing medicines and treatments in the NHS in England and Wales.

This guidance was developed using the NICE [single technology appraisal](#) process.

It replaces NICE technology appraisal guidance 271 (published January 2013). TA271 did not recommend fluocinolone acetonide intravitreal implant for treating chronic diabetic macular oedema considered insufficiently responsive to available therapies. The updated guidance recommends fluocinolone acetonide intravitreal implant as an option for treating chronic diabetic macular oedema that is insufficiently responsive to available therapies only if the implant is to be used in an eye with an intraocular (pseudophakic) lens **and** the manufacturer provides fluocinolone acetonide intravitreal implant with the discount agreed in the patient access scheme.

This guidance has been incorporated into the [NICE pathway on diabetes](#) along with other related guidance and products.

We have produced a [summary of this guidance for patients and carers](#). Tools to help you put the guidance into practice and information about the evidence it is based on are also [available](#).

Your responsibility

This guidance represents the views of NICE and was arrived at after careful consideration of the evidence available. Healthcare professionals are expected to take it fully into account when exercising their clinical judgement. However, the guidance does not override the individual responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or guardian or carer.

Implementation of this guidance is the responsibility of local commissioners and/or providers. Commissioners and providers are reminded that it is their responsibility to implement the guidance, in their local context, in light of their duties to have due regard to the need to eliminate unlawful discrimination, advance equality of opportunity and foster good relations. Nothing in this guidance should be interpreted in a way that would be inconsistent with compliance with those duties.

Copyright

© National Institute for Health and Care Excellence 2013. All rights reserved. NICE copyright material can be downloaded for private research and study, and may be reproduced for educational and not-for-profit purposes. No reproduction by or for commercial organisations, or for commercial purposes, is allowed without the written permission of NICE.

ISBN: 978-1-4731-0368-9

ORIGINAL ARTICLE

Intravitreal Versus Subtenon Triamcinolone Acetonide Injection for Diabetic Macular Edema: A Systematic Review and Meta-analysis

Hui-Ping Qi¹, Sheng Bi^{2,3}, Shu-Qin Wei⁴, Hao Cui¹, and Jing-Bo Zhao⁵

¹Department of Ophthalmology, the First Affiliated Hospital of Harbin Medical University, Harbin, China,

²Department of Neurology, the First Affiliated Hospital of Harbin Medical University, Harbin, China,

³Central Laboratory, the First Affiliated Hospital of Harbin Medical University, Harbin, China,

⁴Perinatal Epidemiology, Sainte-Justine Hospital, University of Montreal, Montreal, Canada, and

⁵Department of Epidemiology, School of Public Health, Harbin Medical University, Harbin, China

ABSTRACT

Objective: To compare the efficacy of intravitreal (IV) triamcinolone acetonide (IVTA) versus subtenon (ST) triamcinolone acetonide (STTA) injection for the treatment of diabetic macular edema (DME).

Methods: Searches for randomized clinical trials published between 1 January 1950 and 15 March 2011 were conducted using PubMed, MEDLINE, EMBASE, and the Cochrane Library included in the present meta-analysis are five randomized controlled trials, each with a minimum follow-up of 3 mo. All included studies evaluated the efficacy of TA for the treatment of refractory DME, and compared IVTA with STTA by measuring visual acuity (VA), central macular thickness (CMT), and intraocular pressure (IOP).

Results: One mo post-injection, treatment with IVTA had significantly improved VA (MD, -0.14 logMAR; 95% CI = -0.16 to -0.13) and reduced CMT (MD = -174.02 μ m; 95% CI = -249.97 to -98.08) compared with STTA. At 3 mo post-injection, treatment with IVTA had significantly improved VA (MD = -0.07 logMAR; 95% CI = -0.09 to -0.05) and reduced CMT (MD = -119.46 μ m; 95% CI = -176.55 to -62.36) compared with STTA. The benefits of either treatment were no longer significant at 6 mo, and patients had to be retreated. Compared with STTA, IVTA injections produced no difference in IOPs at 1 mo, higher IOPs at 3 mo, and lower IOP values at 6 months.

Conclusions: Within 3 mo, IVTA is more effective than is STTA in improving VA and reducing CMT in patients with refractory DME. However, the benefits of either regimen were no longer evident at 6 mo.

KEYWORDS: Diabetic macular edema, Triamcinolone acetonide, Intravitreal, Subtenon, Meta-analysis

INTRODUCTION

Macular edema (ME) is a frequent manifestation of diabetic retinopathy (DR) and the leading cause of visual impairment in patients with diabetes mellitus (DM).^{1–3} Diabetic ME (DME), which is characterized by retinal thickening within 2 disc diameters of the center of the macula,⁴ results from damage to the internal blood-retinal barrier (BRB). Permeability of retinal capillaries is increased, and leakage of fluid

from blood vessels within the macula leads to deterioration of visual acuity (VA). With the prevalence of DM estimated to double during the next 20 years,⁵ it is likely that, unless adequate treatment is provided, DME will be responsible for significant loss of vision in the future.⁶

Reports from the Early Treatment Diabetic Retinopathy Study (ETDRS) indicate that focal grid laser photocoagulation for clinically significant ME effectively reduces the risk of progressive visual

Both Jing-Bo Zhao and Hui-Ping Qi are corresponding authors.

Received 11 August 2011; revised 28 April 2012; accepted 19 June 2012

Correspondence: Jing-Bo Zhao, Department of Epidemiology, School of Public Health, Harbin Medical University, No.157 Baojian Street, Nangang District, Harbin 150086, Heilongjiang Province, People's Republic of China. E-mail: zhaojb168@sina.com; Hui-Ping Qi, Department of Ophthalmology, the First Affiliated Hospital of Harbin Medical University, No. 23 Youzheng Street, Nangang District, Harbin 150001, Heilongjiang Province, People's Republic of China. E-mail: huipingqi@gmail.com

Efficacy and Safety of Intravitreal Therapy in Macular Edema Due to Branch and Central Retinal Vein Occlusion: a Systematic Review

Amelie Pielen^{1,2*}, Nicolas Feltgen³, Christin Isserstedt³, Josep Callizo³, Bernd Junker^{1,2}, Christine Schmucker⁴

1 University Eye Hospital, Albert Ludwig University, Freiburg, Germany, **2** University Eye Hospital, Medical School of Hannover, Hannover, Germany, **3** University Eye Hospital, Georg-August-University, Goettingen, Germany, **4** German Cochrane Centre, Institute of Medical Biometry and Medical Informatics, Department of Medical Biometry and Statistics, University Medical Centre, Freiburg, Freiburg, Germany

Abstract

Background: Intravitreal agents have replaced observation in macular edema in central (CRVO) and grid laser photocoagulation in branch retinal vein occlusion (BRVO). We conducted a systematic review to evaluate efficacy and safety outcomes of intravitreal therapies for macular edema in CRVO and BRVO.

Methods: And Findings: MEDLINE, Embase, and the Cochrane Library were systematically searched for RCTs with no limitations of language and year of publication. 11 RCTs investigating anti-VEGF agents (ranibizumab, bevacizumab, aflibercept) and steroids (triamcinolone, dexamethasone implant) with a minimum follow-up of 1 year were evaluated.

Efficacy: CRVO: Greatest gain in visual acuity after 12 months was observed both under aflibercept 2 mg: +16.2 letters (8.5 injections), and under bevacizumab 1.25 mg: +16.1 letters (8 injections). Ranibizumab 0.5 mg improved vision by +13.9 letters (8.8 injections). Triamcinolone 1 mg and 4 mg stabilized visual acuity at a lower injection frequency (-1.2 letters, 2 injections).

BRVO: Ranibizumab 0.5 mg resulted in a visual acuity gain of +18.3 letters (8.4 injections). The effect of dexamethasone implant was transient after 1.9 implants in both indications.

Safety: Serious ocular adverse events were rare, e.g., endophthalmitis occurred in 0.0-0.9%. Major differences were found in an indirect comparison between steroids and anti-VEGF agents for cataract progression (19.8-35.0% vs. 0.9-7.0%) and in required treatment of increased intraocular pressure (7.0-41.0% vs. none). No major differences were identified in systemic adverse events.

Conclusions: Anti-VEGF agents result in a promising gain of visual acuity, but require a high injection frequency. Dexamethasone implant might be an alternative, but comparison is impaired as the effect is temporary and it has not yet been tested in PRN regimen. The ocular risk profile seems to be favorable for anti-VEGF agents in comparison to steroids. Because comparative data from head-to-head trials are missing currently, clinicians and patients should carefully weigh the benefit-harm ratio.

Citation: Pielen A, Feltgen N, Isserstedt C, Callizo J, Junker B, et al. (2013) Efficacy and Safety of Intravitreal Therapy in Macular Edema Due to Branch and Central Retinal Vein Occlusion: a Systematic Review. PLoS ONE 8(10): e78538. doi:10.1371/journal.pone.0078538

Editor: Demetrios Vavvas, Massachusetts Eye & Ear Infirmary, Harvard Medical School, United States of America

Received: June 3, 2013; **Accepted:** September 16, 2013; **Published:** October 25, 2013

Copyright: © 2013 Pielen et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: These authors have no support or funding to report.

Competing interests: The authors have the following interests. Amelie Pielen is consultant for Allergan, Bayer HealthCare and Novartis Pharma, investigator in multicenter clinical trials sponsored by Bayer HealthCare, Novartis Pharma, GlaxoSmithKline, Pfizer, Alcon, Accovion and Genentech, received research grants from Novartis Pharma and travel grants from Bayer HealthCare. Nicolas Feltgen is consultant for Bayer, Novartis and Allergan, investigator in clinical trials sponsored by Bayer, Novartis and Allergan and received research and travel grants from Novartis. Josep Callizo is consultant for Bayer, Allergan and Novartis, investigator in clinical trials sponsored by Bayer, Allergan and Novartis and received travel grant from Novartis. Bernd Junker is investigator in multicenter clinical trials sponsored by Bayer HealthCare, Novartis Pharma, GlaxoSmithKline, Pfizer, Alcon, Accovion and Genentech and functioned as lecturer for Heidelberg Engineering, Allergan and Novartis. This does not alter their adherence to all the PLOS ONE policies on sharing data and materials.

* E-mail: amelie.pielen@uniklinik-freiburg.de

Introduction

Macular edema is the main cause of visual impairment in central retinal vein occlusion (CRVO) and branch retinal vein occlusion (BRVO) [1,2]. In recent years, anti-inflammatory and anti-angiogenic therapeutic strategies have been used to target vascular permeability and leakage to reduce macular edema and improve vision.

Among corticosteroids, triamcinolone acetonide, dexamethasone and fluocinolone have shown potential to reduce edema in RVO [3–6]. However, their drawbacks are known side effects, such as cataract progression and rise of intraocular pressure (IOP). The off-label triamcinolone acetonide is crystalline and is commercially available as Kenalog (Kenalog, 40 mg/mL, Bristol-Myers Squibb, Princeton, NJ), Trivaris (Pharm Allergan Inc., Irvine California) or prepared for injection with methocel [7]. Dexamethasone is more potent and soluble in comparison to triamcinolone and various attempts have been made to construct a slow-release device. Ozurdex™ (Pharm Allergan Inc., Irvine California) was the first dexamethasone implant that was approved by the Food and Drug Administration (FDA) for intraocular use in chronic uveitis and macular edema due to RVO in June 2009. Ozurdex shows an anti-edematous effect up to 6 months [8,9]. A fluocinolone implant is currently under investigation that is supposed to even last for up to 3 years [6].

Beside corticosteroids, progress in anti-angiogenic drug development provides us with multiple new therapeutic agents based on a concept of modified antibodies versus vascular endothelial growth factor (VEGF) and related molecules. For example, pegaptanib sodium, a 40-kDa RNA aptamer, binds to the isoform 165 of VEGF and was first approved for intravitreal therapy in wet age-related macular degeneration (AMD) [10,11]. A possible effect on visual acuity in macular edema due to CRVO and BRVO was investigated but pegaptanib did not receive approval for these indications [12,13]. Currently, the most frequently used anti-VEGF agents are ranibizumab (Lucentis™, Genentech, Inc., South San Francisco, CA, and Novartis Pharma AG, Basel, Switzerland) and bevacizumab (Avastin™, Genentech, Inc., South San Francisco, CA, and Roche, Basel, Switzerland). Ranibizumab is an antibody fragment with a high binding affinity towards all forms of VEGF. It has been approved for intravitreal therapy of wet AMD, diabetic macular edema and RVO by the FDA and European Medicines Agency (EMA) between 2006 and 2012 (depending on the indication). Bevacizumab is widely used as off-label intravitreal therapy. The cost difference to ranibizumab is striking. Large Head-to-head studies sponsored by the American and British public authorities have been conducted to compare both drugs in wet AMD (CATT [14] and IVAN [15], respectively). One and two year data of these trials showed no inferiority. Despite the approval of ranibizumab for macular edema in RVO, we assume that off-label bevacizumab is used as frequently in RVO as in AMD. This assumption is supported by a vast number of publications on the use of bevacizumab for RVO in clinical settings (e.g. [16–18]). The latest developed anti-angiogenic drug is aflibercept (VEGF-trap eye, Eylea™, Regeneron Pharmaceuticals, Inc., and Bayer Pharma AG,

Berlin, Germany), a 115-kDa decoy receptor fusion protein, composed of the second domain of human VEGF receptor 1 and the third domain of VEGF receptor 2, which are fused to the Fc domain of human IgG1 [19,20]. A recent publication reported that the binding affinity for this drug is even higher than for ranibizumab and bevacizumab, respectively [21]. Aflibercept was first approved for intravitreal therapy of wet AMD in 2011. FDA approval for RVO was granted in September 2012. EMA approval is expected for 2013.

Most of this pharmaceutical research was conducted at the same time. Notably, the multicenter randomized controlled trials (RCTs) leading to FDA approval in RVO were conducted in parallel, which resulted in homogenous control groups using observation and/or sham injection in CRVO and sham injection and/or grid laser photocoagulation in BRVO [9,20,22,23]. These research proceedings, however, resulted in a lack of head-to-head studies, comparing the various drugs to each other. Currently the ophthalmologist can choose between different therapeutic options to treat vision loss due to macular edema in RVO. Some of the therapeutic options are approved, but others are used off-label. To support the ophthalmologist's decision between the different therapeutic options and to calculate the accompanying risks, we conducted a systematic review to investigate and compare the efficacy and safety of all therapeutic agents currently used in intravitreal therapy in RVO with a follow-up of at least 12 months.

Methods

Search Strategy

We searched Medline, Premedline, Embase and the Cochrane library from inception until November 2012. The search strategy was based on combinations of medical subject headings and keywords and was not restricted to specific languages or years of publications. The search strategy used in Medline is presented in Text S1. Search strategies for other databases were modified to meet the requirements of each database. The searches were supplemented by hand searching the bibliographies of included studies. Currently conducted RCT's on intravitreal treatment of macular edema in RVO were searched in the register for clinical trials (<http://clinicaltrials.gov>) and in the WHO International Clinical Trial Registry Platform (<http://www.who.int/ictrp/en/>).

Inclusion and exclusion criteria

We included RCTs which investigated intravitreal ranibizumab, bevacizumab, pegaptanib, aflibercept, triamcinolone, dexamethasone, or fluocinolone as monotherapy in comparison to any non-surgical control group (observation or sham injection in CRVO, sham injection and/or grid laser photocoagulation in BRVO). Trials which compared one substance to another or to grid laser photocoagulation were also included. To address long-term efficacy and systematic adverse effects, one year follow-up data (or longer) had to be available.

Studies were excluded if participants received a combination of intravitreal pharmacological agents within the same eye (e.g. ranibizumab and triamcinolone), if both eyes of the same

person were treated as study eyes, and if less than 10 patients per arm were enrolled. Studies which treated patients who had already been treated within 3 months using one of the substances were excluded. To account for the 2 different indications, study results had to be attributable to either BRVO or CRVO separately.

Data Extraction and Quality Assessment

Titles and abstracts were reviewed by using the above mentioned selection criteria. Full papers of appropriate studies were obtained for detailed evaluation. Data extraction and quality assessment were carried out after a modified evaluation tool of the Center for Reviews and Dissemination [24]. Data extraction of *study characteristics* comprised the following data: number of included patients, number of treated patients in each group, control treatment, follow-up time, treatment regimen and dosages, baseline visual acuity, and time between symptoms and/ or diagnosis of macular edema and treatment in months. Additionally, the percentage of patients included within 3 months of disease onset was noted, if reported. If duration was given in weeks or days, values were transformed to mean time in months. Visual acuity is given in number of letters (according to ETDRS scores) and was changed to letters, if reported in logMAR. Data extraction of *study outcome* comprised, if available, the following data: number of included patients, mean number of intravitreal injections within 12 months (or from month 12 to 24), percentage of patients with gain or loss of visual acuity ≥ 15 letters, and mean change in visual acuity from baseline to month 12 (from month 12 to 24) in letters. Despite the multitude of information provided, we extracted information on the percentage (and/or number of patients) of endophthalmitis, uveitis, retinal detachment, retinal tear, and traumatic lens damage, rise of intraocular pressure (IOP), medical treatment of IOP rise, glaucoma surgery, cataract progression and surgery, and vitreous hemorrhage. Given the underlying disease of BRVO or CRVO, we collected also all information on vitreous hemorrhage, neovascularization of the iris (NVI) and neovascular glaucoma (NVG), if reported. Regarding systemic adverse events we extracted information on death (any cause), myocardial infarction, cerebrovascular accidents, nonocular hemorrhage and infection if obtainable. *Parameters to assess the methodological quality* of RCTs and estimate risk of bias were information on blinding, transparency of patient flow, sample size calculation, intention-to-treat analysis, definition of expected adverse events, the method used to collect adverse events data, and comparability between groups at inclusion. All stages of study selection, data extraction and quality assessment were done independently by two reviewers (AP, CI or NF). Any disagreement was resolved by discussion and consensus.

Statistical Analysis

A narrative summary is provided because the identified studies are enormous heterogeneous in terms of intervention and sample size and, therefore, unsuitable for pooling.

Results

Study selection/ Included studies

After removing duplicate references, the searches identified 3374 citations (Medline (1035), Premedline (44), Medline in process (45), Cochrane Library (81) und Embase (2169) (Figure 1) modified from [25]). Of the 3374 citations, 7 RCTs were included to evaluate BRVO and 5 RCTs were included for CRVO.

Study characteristics (Table 1)

Inclusion criteria were met by 7 multicenter RCTs and 4 single-center RCTs investigating CRVO alone ($n = 4$, [5,20,22,26]), BRVO alone ($n = 5$, [4,23,27–29]) or both CRVO/BRVO ($n = 2$, [8,30]). Table 1 presents study characteristics of all identified RCTs evaluating monotherapy of intravitreal triamcinolone, dexamethasone, ranibizumab, bevacizumab or aflibercept in macular edema due to CRVO or BRVO separately with a follow-up of at least 1 year. Control treatment consisted of sham injections or observation in CRVO patients (SCORE-CRVO, GENEVA, CRUISE, Epstein, and COPERNICUS [5,8,20,22,26]). In BRVO control treatment differed among the studies between sham treatment (GENEVA, BRAVO [8,23]), grid laser photocoagulation (SCORE, Parodi, Russo [4,21,22]), and grid laser photocoagulation with bevacizumab (Donati [29]).

Triamcinolone - The SCORE study (CRVO/ BRVO) and Parodi 2008 (BRVO). The SCORE (Standard Care vs. Corticosteroid for Retinal Vein Occlusion) study compared intravitreal triamcinolone (Trivaris, Allergan) 1 mg or 4 mg with observation in CRVO [5] and with grid laser photocoagulation in BRVO [4]. In total, 271 patients were included in SCORE-CRVO and longest follow-up was 36 months. In the parallel SCORE-BRVO trial 411 patients were enrolled [4]. Treatment regimen allowed re-treatment with triamcinolone at PRN criteria using 4 months intervals [4,5]. The mean time from symptoms (or diagnosis) of CRVO to treatment initiation was 4 months, only 39% of the patients were treated within 3 months and 81% within 6 months of onset [5]. Correspondingly, the time span to treatment initiation in SCORE-BRVO was 4 months [4].

Parodi et al. investigated subthreshold grid laser photocoagulation (SGLT) in combination with triamcinolone 4 mg (Kenalog, 40 mg/mL, Bristol-Myers Squibb, Princeton, NJ) versus SGLT and sham injection in 24 patients with BRVO [27]. Follow-up was 12 months. Mean time from symptoms to treatment was 8 months (range 3–18) [27].

Dexamethasone Implant - The GENEVA study (CRVO/ BRVO). The GENEVA (Global Evaluation of Implantable Dexamethasone in Retinal Vein Occlusion With Macular Edema) study evaluated the safety and efficacy of an intravitreal dexamethasone implant 0.7 mg versus an implant 0.35 mg versus sham injection in both BRVO and CRVO. The primary outcome measure was time to visual improvement of 15 letters or more. This outcome was pooled for BRVO and CRVO patients [8]. GENEVA included 1256 patients in their open-label 6-month extension of which 1131 completed the trial. In total, 34.5% of patients presented with CRVO and 65.5% with BRVO in the GENEVA study. Time from symptoms

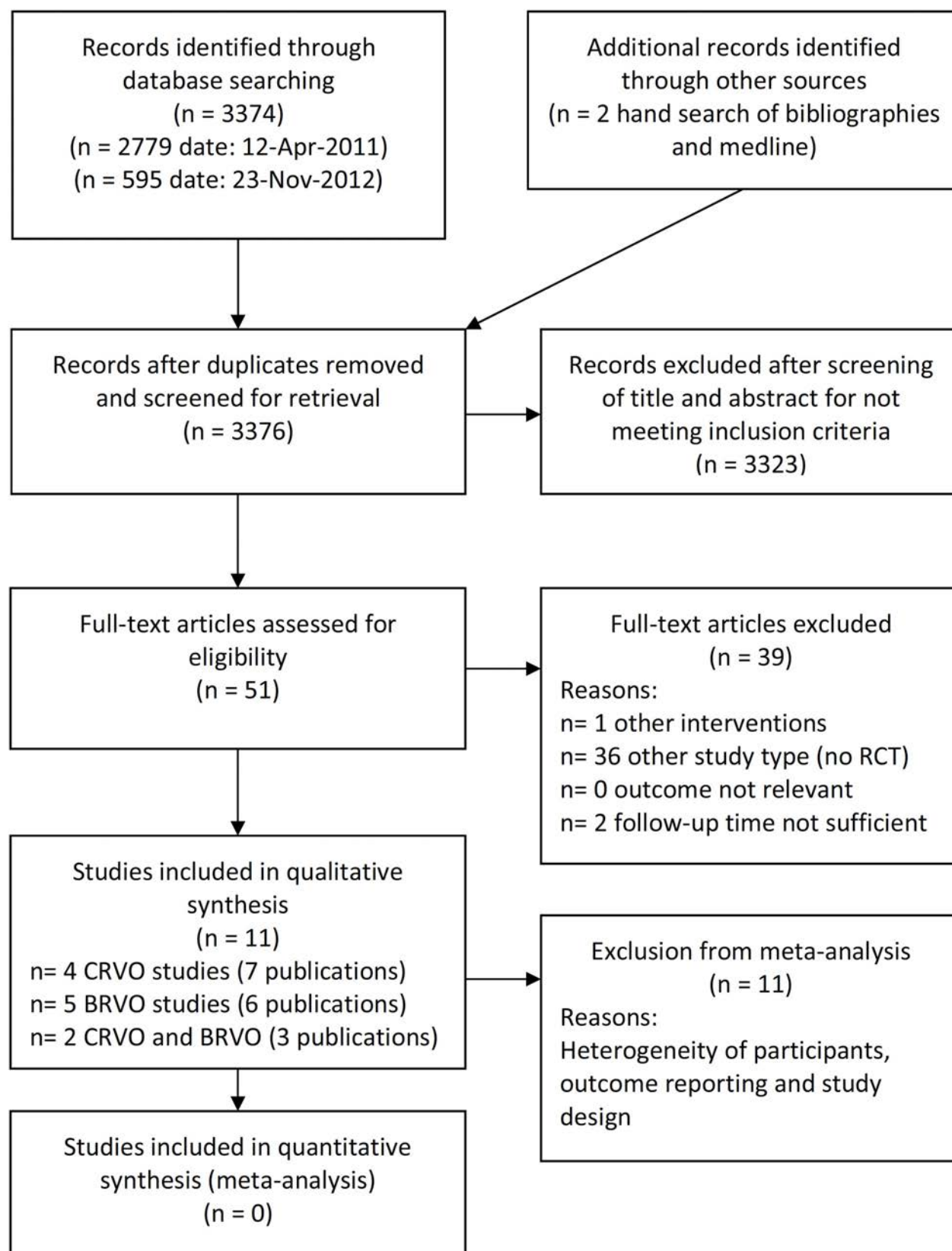


Figure 1. Flow chart of literature search and study selection.

doi: 10.1371/journal.pone.0078538.g001

Table 1. Study characteristics.

Study	Investigational product	Included patients	Treated patients			Follow-up (months)	Treatment regime	Dosage (mg)*	Time between symptoms/ diagnosis and treatment (months)			Baseline BCVA (letters)		
			1 mg	4 mg	Control treatment				1 mg	4 mg	Obs.	1 mg	4 mg	Obs.
CRVO	SCORE 2009, 2011[5,32]	Triamcinolone at 3 years)	271 (81											
			91	91	88	12 (36)	PRN every 4 months	1 and 4 (Trivaris)	4.5	4.2	4.2	51	51	52
			0.7 mg	0.35 mg	Sham/ 0.7 mg									
CRVO/ BRVO	GENEVA 2011 [8,9]	Dexamethasone implant	1256	421	412	423	12	every 6 months	0.35 and 0.7 implant/ Sham injection	5.2 (15-18% < 3)		54		
	CRVO#		31.9%	37.2%	34.5%									
	BRVO#		68.1%	62.8%	65.5%									
			1.25 mg		Sham/ 1.25 mg				1.25 mg	Sham/ 1.25 mg		1.25 mg	Sham/ 1.25 mg	
CRVO	Epstein 2012 [26]	Bevacizumab	60	30	30	12	1.25 mg/ 6 weeks or Sham/ 6 weeks + 1.25 mg/ 6 weeks	1.25	1.9	2.2		44.4	43.9	
			2 mg		Sham/ 2 mg				2 mg	Sham/ 2 mg		2 mg	Sham/ 2 mg	
CRVO	COPERNICUS 2013 [20,36]	Aflibercept	189	115	74	12	Monthly for 6 months then PRN	2	2.73 (56.1% ≤2)	1.88 (71.2% ≤2)		50.7	48.9	
			0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg									
CRVO	CRUISE 2011 [22]	Ranibizumab	392	132	130	130	12	Monthly for 6 months then PRN	0.3 and 0.5	3.3 (69% ≤ 3)		48.3		
CRVO/ BRVO	HORIZON 2012 [30]	Ranibizumab	181 of 304	70 of 107	51 of 99	60 of 98	12 - 24	PRN every 1-3 months	0.5	(idem CRUISE)		(idem CRUISE)		
			0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg (rescue: grid)									
BRVO	BRAVO 2011 [23]	Ranibizumab	397	134	131	132	12	Monthly for 6 months then PRN	0.3 and 0.5	3.5 (65% ≤ 3)		54.6		
BRVO/ CRVO	HORIZON 2012 [30]	Ranibizumab	205 of 304	66 of 103	73 of 104	66 of 97	12- 24	PRN every 1-3 months	0.5	(idem BRAVO)		(idem BRAVO)		
			1 mg	4 mg	Grid				1 mg	4 mg	Grid			
BRVO	SCORE 2009, 2011 [4,32]	Triamcinolone	411 (128 at 3 years)	136	138	137	12 (36)	PRN every 4 months	1 and 4 (Trivaris)	4.5 (38% ≤ 3)	4.6 (37% ≤ 3)	4.5 (36% ≤ 3)	57	
PLOS ONE www.plosone.org			SGLT + 4 mg			SGLT + 5 sham			October 2013 Volume 8 Issue 10 e78538					

Table 1 (continued).

Study	Investigational product	Included patients	Treated patients	Control treatment	Follow-up (months)	Treatment regime	Dosage (mg)*	Time between symptoms/ diagnosis and treatment (months)	Baseline BCVA (letters)
BRVO Parodi 2008 [27]	Triamcinolone	24	11	13	12	4 mg triamcinolone + SGLT after 2 weeks or SGLT + sham injection	4 in 0.1 ml (Kenalog)	8	Nr
			1.25 mg	Grid					1.25 mg Grid
BRVO Russo 2009 [28]	Bevacizumab	30	15	15	12	PRN	1.25	Nr	41.5 40.5
			1.25 mg	1.25 mg + grid					1.25 mg 1.25 mg + grid
BRVO Donati 2012 [29]	Bevacizumab	18	9	9	12	Monthly for 3 then PRN at month 3 and 6	1.25	3.17	50 55

Nr: Not reported, PRN: pro re nata, Obs.: Observation, SGLT: subthreshold grid laser photocoagulation

* Dosage in mg per 0.05 ml volume if not indicated separately

Percentage of CRVO/ BRVO patients at baseline of core-study [9]

doi: 10.1371/journal.pone.0078538.t001

to treatment was 5.2 months, 15-18% of patients received treatment within 3 months of diagnosis.

Ranibizumab - The CRUISE (CRVO) and BRAVO (BRVO) studies. The CRUISE (Ranibizumab for the Treatment of Macular Edema After Central Retinal Vein Occlusion Study: Evaluation of Efficacy and Safety) study enrolled 392 patients and assigned them either to ranibizumab 0.3 mg or 0.5 mg monthly for 6 months followed by PRN treatment of ranibizumab 0.5 mg in all groups from month 6 to 12 [22]. Control treatment in this study consisted of sham injections until month 6. Mean duration of symptoms to treatment was 3.3 months. In contrast to SCORE, the percentage of patients receiving treatment within 3 months of CRVO onset was considerably higher (69% in CRUISE versus 39% in SCORE [5,22]).

The BRAVO (Ranibizumab for the Treatment of Macular Edema Following Branch Retinal Vein Occlusion: Evaluation of Efficacy and Safety) study assigned 397 patients to 3 treatment arms (corresponding to CRUISE): intravitreal ranibizumab 0.3 mg, intravitreal ranibizumab 0.5 mg and sham injections for 6 months followed by a PRN of ranibizumab 0.5 mg for all patients until month 12 [23]. Grid laser photocoagulation was only allowed as rescue therapy (to all groups), but not earlier than 3 months after inclusion [23]. 65% of BRVO patients were treated within 3 months of RVO duration, mean time from symptoms to treatment was 3.5 months [23].

Ranibizumab in the 2nd year - The HORIZON trial (BRVO and CRVO). In cohort 2 of the open-label, single-arm, multicenter extension HORIZON trial, patients from BRAVO and CRUISE studies were included [30]. 304 patients of each trial continued treatment of macular edema with intravitreal ranibizumab 0.5 mg at PRN intervals (between 1 and 3 months) for additional 12 months. Grid laser photocoagulation was eligible as rescue therapy for BRVO patients. The

HORIZON trial was discontinued after FDA approval for ranibizumab in RVO (in accordance to the study protocol). Finally, 181 of 304 CRVO patients and 205 of 304 BRVO patients were analysed. Regarding the safety analysis, data of all patients over the entire study duration of 24 months were included [30].

Bevacizumab – Epstein 2012 (CRVO), Russo 2009 and Donati 2012 (BRVO). Intravitreal bevacizumab in CRVO was evaluated in the RCT by Epstein and colleagues [26,31]. Patients received either intravitreal bevacizumab 1.25 mg or sham injection every 6 weeks over a period of 6 months. After this, an open label extension followed during which all patients received bevacizumab 1.25 mg every 6 weeks. Each group comprised 30 patients, who showed symptoms of CRVO for 1.9 months in the bevacizumab and 2.2 months in the sham/bevacizumab group.

Russo et al. compared intravitreal bevacizumab 1.25 mg after PRN schemata versus grid laser photocoagulation in 30 patients with BRVO over 12 months [28]. Inclusion criterion was duration of macular edema of at least 3 months.

Donati et al. randomized 18 patients (1:1) to 3 intravitreal injections of bevacizumab 1.25 mg or intravitreal injections combined with grid laser photocoagulation (one week after first injection) [29]. Patients could receive further injections PRN. Follow-up visits were conducted at month 3, 6 and 12 (final visit). Thus the maximum number of injections per group was 5. Median duration of macular edema before treatment was 3.2 and 3.1 months, respectively [29].

Aflibercept - The COPERNICUS trial (CRVO). Only recently, aflibercept has been investigated in patients with CRVO in the COPERNICUS (Controlled Phase 3 Evaluation of Repeated Intravitreal Administration of VEGF Trap-Eye in Central Retinal Vein Occlusion: Utility and Safety) trial [20]. 189 patients with CRVO were included and received intravitreal

afibercept 2 mg (afibercept group) compared to sham treatment (sham/afibercept group) at a randomisation ratio of 2:1 during the first 6 months followed by (monthly) PRN treatment of intravitreal afibercept 2 mg in all patients. Primary endpoint was the proportion of patients with an improvement of BCVA ≥ 15 letters. Mean time from symptoms to treatment was relatively short with 2.7 months (afibercept) and 1.9 months (sham/afibercept). First treatment within 2 months occurred in 56.1% (afibercept) and 71.2% (sham/afibercept), respectively [20].

Study results: Efficacy Outcome (Table 2, Figure 2, 3)

Systematic extraction of outcome measures was limited by different primary outcome measures reported in the studies.

Triamcinolone - The SCORE study (CRVO/ BRVO) and Parodi 2008 (BRVO). The SCORE-CRVO study found a significant difference between triamcinolone and sham: The percentage of participants with a gain of visual acuity of 15 letters or more was 26.5%, 25.6% and 6.8% for triamcinolone 1 mg, 4 mg and sham, respectively ($p=0.001$ for 1 mg versus observation and also for 4 mg versus observation) [5]. Both triamcinolone concentrations stabilized visual acuity at month 12 (-1.2 letters), whereas sham injections led to a mean deterioration of -12.1 letters [5]. The percentage of participants who lost 15 letters or more was nearly doubled in the sham group (44% at month 12) compared to 25% (1 mg) and 26% (4 mg) in the triamcinolone groups [5]. This remained almost unchanged until 24 months (48% [sham] versus 31% [1 mg] and 26% [4 mg]). Mean number of injections was 2.2 (triamcinolone 1 mg), 2.0 (triamcinolone 4 mg) and 0.1 (observation).

The SCORE-BRVO study did not find a statistically significant difference of visual acuity gain after 12 months: 26% (triamcinolone 1 mg), 27% (4 mg) and 29% (grid laser photocoagulation) [4]. Mean change in BCVA was +4 to +5 letters in all groups. It was achieved with a mean treatment application of 2.2 injections (for triamcinolone 1 mg), 2.1 injections (for triamcinolone 4 mg), and 1.8 (for grid laser photocoagulations, if no dense macular hemorrhages were present at baseline; 0.7 applications for the group, that received deferred laser after resorption of macular hemorrhages) [4].

Parodi et al. found a visual acuity gain of more than 15 letters in 54% of the combination group (SGLT and triamcinolone) and 38% in controls (SGLT with sham injection) after a single treatment [27]. Mean visual acuity increased by +17 letters (intervention group) and +6.5 letters (control group), respectively ($p=0.011$ for mean change in visual acuity at month 12 compared to baseline in the combination group).

Dexamethasone Implant - The GENEVA study (CRVO/ BRVO). Primary outcome of the GENEVA open-label extension study until month 12 was safety [8]. Efficacy within the core study was measured in time to achieve a significant improvement (gain of more than 15 letters). Patients treated with a dexamethasone implant reached this goal significantly faster and in a higher percentage than sham treated patients (pooled data from BRVO and CRVO, $p<0.001$) [8,9]. Looking into BRVO and CRVO data separately, the GENEVA study did

not find a significant difference in visual acuity at month 12 between the groups. Unfortunately, numbers are not provided within the publication, but figures indicate that change in BCVA was +2 letters in dexamethasone 0.7 mg/0.7 mg versus -1 letter in sham/0.7 mg in CRVO and +6 letters in BRVO in both groups [8]. Mean number of injections was 1.9 (0.7 mg/0.7 mg), 1.9 (0.35 mg/0.7 mg), and 0.8 (sham/0.7 mg).

Ranibizumab - The CRUISE (CRVO), BRAVO (BRVO) and HORIZON trial (CRVO/ BRVO). The CRUISE study selected the mean change from baseline BCVA letter score at month 12 as primary outcome, and also reported all parameters chosen for extraction within this search [22]. Patients treated with ranibizumab gained +13.9 letters (both groups 0.3/0.5 mg and 0.5/0.5 mg independently), and +7.3 letters, if treated 6 months with sham injections followed by ranibizumab 0.5 mg ($p<0.001$ for each ranibizumab group versus sham/0.5 mg group) [22]. Percentage of patients with a gain of 15 letters or more was highest in ranibizumab 0.5/0.5 mg (50.8%), versus 0.3/0.5 mg (47.0%) and sham/0.5 mg (33.1%). These gains were achieved by 9.6 (0.3/0.5 mg), 8.8 (0.5/0.5 mg), and 9.1 (sham/0.5 group [5.4 sham injections and 3.7 IVR injections]) injections after month 12 [22].

Results of the BRAVO trial found a mean change in BCVA of +16.4 letters (0.3/0.5 mg), +18.3 letters (0.5/0.5 mg) and +12.1 (sham/0.5 mg) ($p<0.01$ for each ranibizumab group versus sham/0.5 mg group) [27]. More patients treated with ranibizumab from the beginning gained ≥ 15 letters (56% and 60.3%) than those with sham injections followed by ranibizumab 0.5 mg (43.9%). Mean number of injections did not differ between groups (mean 8.3, 8.4 and 9.4 injections, respectively).

For patients from BRAVO and CRUISE studies who continued from month 12 to month 24 within the HORIZON trial visual acuity decreased slightly (in CRVO) or remained stable (in BRVO) [30]. In HORIZON-CRVO mean visual acuity change was -5.2 (ranibizumab 0.3/0.5 mg), -4.1 (0.5/0.5 mg) and -4.2 (sham/0.5 mg) letters, respectively. Injections were given significantly less frequently during the second year, the mean number of injections was 3.8 (0.3/0.5 mg), 3.5 (0.5/0.5 mg) and 2.9 (sham/0.5 mg). The mean total change in BCVA from baseline to month 24 resulted in +8.2 (0.3/0.5 mg), +12 (0.5/0.5 mg), and +7.6 (sham/0.5 mg) letters.

HORIZON-BRVO: Patients with BRVO stabilized at the BCVA letter score gained at month 12. Mean BCVA change during the second year was -2.3 letters (0.3/0.5 mg), -0.7 letter (0.5/0.5 mg) and +0.9 letter (sham/0.5 mg). The mean difference in BCVA score between baseline and month 24 was +14.9 letters, +17.5 letters and +15.6 letters, respectively. Frequency of intravitreal injections was considerably lower than in the first year and even lower than in CRVO patients during year 2. Patients received 2.4 (0.3/0.5 mg), 2.1 (0.5/0.5 mg) and 2.0 (sham/0.5 mg) injections [30]. Caution: Numbers are probably that low because the HORIZON trial was terminated immediately after FDA approval of ranibizumab.

Bevacizumab - Epstein 2012 (CRVO), Russo 2009 and Donati 2012 (BRVO). Epstein and colleagues found that CRVO patients treated with bevacizumab 1.25 mg every 6 weeks from baseline gained +16.1 letters, compared to +4.6

Table 2. Efficacy parameters.

Study	Investigational product	Included patients	Injections per patient (mean number/ 12 months)			Percentage of patients with BCVA gain ≥ 15 letters (at 12 months)			Change in BCVA at month 12 (mean number of letters)			Percentage of patients with BCVA loss ≥ 15 letters		
			1 mg	4 mg	Observation	1 mg	4 mg	Observation	1 mg	4 mg	Observation	1 mg	4 mg	Observation
CRVO	SCORE 2009, 2011 [5,32]	271	2.2	2.0	0.1	26.5 %	25.6 %	6.8 %	-1.2	-1.2	-12.1	25% (24 mo)	26% (26 mo)	44% (48% 24 mo)
			0.7/ 0.7 mg											
CRVO/ BRVO	GENEVA 2011 [8,9]	1256#	1.86	1.85	0.83	Nr			Nr			Nr		
Subgroup CRVO	GENEVA 2011 [8]	437 (at baseline)	Nr			Nr			0.7/ 0.7 mg	Sham/ 0.7 mg		Nr		
Subgroup BRVO	GENEVA 2011 [8]	830 (at baseline)	Nr			Nr			0.7/ 0.7 mg	Sham/ 0.7 mg		Nr		
CRVO	Epstein 2012 [26]	60	8	4/ 4		60%	33.3%		1.25 mg	Sham/ 1.25 mg	1.25 mg	1.25 mg	Sham/ 1.25 mg	Sham/ 1.25 mg
CRVO	COPERNICUS 2013 [20]	189	8.5	Sham/ 2 mg	5.3 / 3.9	55.3%	30.1%		2 mg	Sham/ 2 mg	2 mg	2 mg	Sham/ 2 mg	Sham/ 2 mg
CRVO	CRUISE 2011 [22]	392	9.6	8.8	5.4/ 3.7	47%	50.8%	33.1%	0.3/ 0.5 mg	Sham/ 0.5 mg	0.3/ 0.5 mg	0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg
CRVO	HORIZON 2012 [30]	181/ 304	3.8 (12-24)	3.5 (12-24)	2.9 (12-24)	38.6% (from baseline)	45.1%	38.3%	-5.2 (12-24), +8.2 (0-24)	-4.1 (12-24), +12 (0-24)	-4.1 (12-24), +7.6 (0-24)	12.9%	5.9%	13.3%
BRVO	BRAVO 2011 [23]	397	8.3	8.4	5.7/ 3.6	56%	60.3%	43.9%	+16.4	+18.3	+12.1	0%	1.5%	4.5%
BRVO	HORIZON 2012 [30]	205/ 304	2.4 (12-24)	2.1 (12-24)	2	50% (from baseline)	60.3%	51.5%	-2.3 (12-24), +14.9 (0-24)	-0.7 (12-24), +17.5 (0-24)	+0.9 (12-24), +15.6 (0-24)	1.5%	1.4%	1.5%
BRVO	SCORE 2009, 2011 [4,32]	411 (128)	2.2	2.1	0.7 (+ H)*, 1.8 (- H)*	26%	27%	29%	+5.7	+4	+4.2	12%	12%	15%
			4 mg + SGLT	SGLT	SGLT	4 mg + SGLT	SGLT	SGLT	4 mg + SGLT	SGLT	SGLT	4 mg + SGLT	SGLT	SGLT

Table 2 (continued).

Study	Investigational product	Included patients	Injections per patient (mean number/ 12 months)	Percentage of patients with BCVA gain ≥ 15 letters (at 12 months)	Change in BCVA at month 12 (mean number of letters)	Percentage of patients with BCVA loss ≥ 15 letters
BRVO Parodi 2008 [27]	Triamcinolone	24	1	54%	+17	0%
			1.25 mg	Grid	+6.5	15%
BRVO Russo 2009 [28]	Bevacizumab	30	1.7	1.25 mg	Grid	1.25 mg
			1.5	80%	+15.5	Grid
BRVO Donati 2012 [29]	Bevacizumab	18	1.25 mg	1.25 mg + grid	1.25 mg	Nr
			4 (median)	33.3%	+15	1.25 mg + grid
BRVO				77.8%	+20	0%

Nr: Not reported, SGLT: subthreshold grid lasercoagulation
* SCORE definition of two groups within the grid treatment arm: (+ H) = dense macular hemorrhages at baseline, (- H) = no hemorrhages at baseline
Patients per treatment groups in the Geneva trial: 330 0.7/ 0.7 mg, 316 0.35/ 0.7 mg, 313 sham/ 0.7 mg, 53 0.7 mg/ none, 57 0.35 mg/ none, 62 sham/ none
doi: 10.1371/journal.pone.0078538.t002

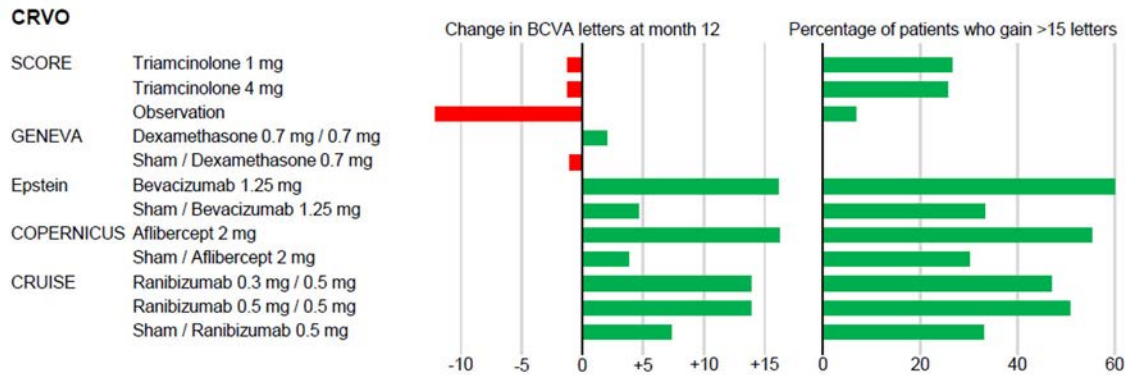


Figure 2. Bar graph for results in central retinal vein occlusion (CRVO). Change in best corrected visual acuity (BCVA) in letters at month 12 for each treatment group (left) and percentage of patients with gain of BCVA of equal or more than 15 letters (right).

doi: 10.1371/journal.pone.0078538.g002

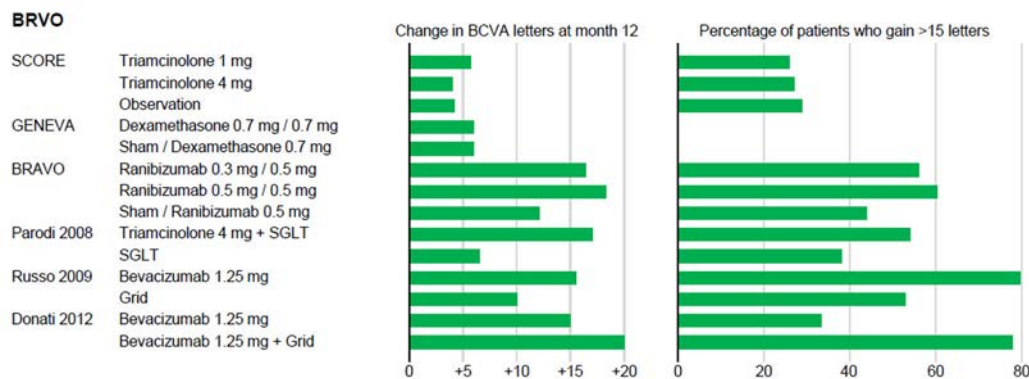


Figure 3. Bar graph for results in branch retinal vein occlusion (BRVO). Change in best corrected visual acuity (BCVA) in letters at month 12 for each treatment group (left) and percentage of patients with gain of BCVA of equal or more than 15 letters (right); SGLT: subthreshold grid laser photocoagulation.

doi: 10.1371/journal.pone.0078538.g003

letters in those treated with sham injections followed by bevacizumab ($p<0.05$) [26]. The percentage of BCVA gain of more than 15 letters was 60% (bevacizumab) vs. 33.3% (sham/bevacizumab), respectively ($p<0.05$).

Russo et al. found a change in BCVA of +15.5 letters in bevacizumab treated patients and +10 letters in grid laser photocoagulation ($p<0.05$), achieved by a mean number of 1.7 intravitreal injections of bevacizumab and 1.5 grid applications [28].

Donati et al. found an improvement in BCVA of +15 letters in patients treated with bevacizumab alone (median: 4 injections) compared to +20 letters in patients who received a combination of intravitreal bevacizumab (median 3 injections) and one grid laser photocoagulation ($p<0.01$ for visual acuity at month 12 versus baseline in each group, $p=0.03$ for difference between groups at month 12) [29]. Three of 9 bevacizumab patients gained more than 15 letters, in the group of patients who received combination therapy 7 of 9 patients reached this gain.

Aflibercept - The COPERNICUS trial (CRVO). CRVO patients in the aflibercept 2 mg group gained +16.2 letters at month 12, compared to +3.8 letters in the sham/aflibercept ($p<0.001$) [20]. The percentage of patients with a gain of more than 15 letters was 55.3% (aflibercept) and 30.1% (sham/aflibercept), respectively ($p<0.001$). After 5.8 aflibercept injections until month 6, visual acuity was stabilized with an additional 2.7 injections during PRN. In the control group, patients with a mean of 5.3 sham injections in the beginning received a mean of 3.9 injections of aflibercept 2 mg until month 12 [20].

Study results: Safety Outcome – ocular adverse events (Table 3, 4)

Triamcinolone - The SCORE study (CRVO/ BRVO) and Parodi 2008 (BRVO). The SCORE CRVO found no significantly increased rate of endophthalmitis, uveitis and retinal detachment after a total of 586 intravitreal triamcinolone injections at 12 months [5]. However, IOP rise and cataract

Table 3. Ocular safety – Endophthalmitis, uveitis, retinal detachment, retinal tear, traumatic lens damage.

		Investigational product	Total number of injections	Endophthalmitis [%], (number of eyes)			Uveitis [%], (number of eyes)			Retinal detachment [%], (number of eyes)			Retinal tear [%], (number of eyes)			Traumatic lens damage [%], (number of eyes)
Study																
CRVO	SCORE 2009, 2011 [5,32]	Triamcinolone	586	0			0			0			Nr			Nr
CRVO/BRVO	GENEVA 2011 [8,9]	Dexamethasone implant	1715 implants [#]	0			Nr			Rare			Rare			Nr
CRVO	Epstein 2012 [26,31]	Bevacizumab	720	0			Nr			0			0			Nr
				2 mg			Sham/ 2 mg			2 mg			Sham/ 2 mg			
CRVO	COPERNICUS 2013 [20,36]	Aflibercept	1266 [*]	0.9% (1)	0		Nr			Nr			0		1.7% (1)	Nr
				0.3/0.5 mg	0.5/0.5 mg	Sham/0.5 mg	0.3/0.5 mg	0.5/0.5 mg	Sham/0.5 mg	0.3/0.5 mg	0.5/0.5 mg	Sham/0.5 mg	0.3/0.5 mg	0.5/0.5 mg	Sham/0.5 mg	
CRVO	CRUISE 2011 [22]	Ranibizumab	2892 [*]	0			2.3% (3)	1.6% (2)	3.9% (5)/1.8% (2)	0			0		1.6% (2)	0/1.8% (2)
CRVO/BRVO	HORIZON 2012 [30]	Ranibizumab		0.9% (2)	0	0	0			Nr			0			Nr
BRVO	BRAVO 2011 [23]	Ranibizumab	2679 [*]	0	0.8% (1)	0	2.2% (3)	0	3.1% (4)/0.9% (1)	0.7% (1)	0	0	0.7% (1)	0	0	0
BRVO/CRVO	HORIZON 2012 [30]	Ranibizumab		0			0			Nr			0			Nr
				1 mg			4 mg			Grid			1 mg			
BRVO	SCORE 2009, 2011 [5,32]	Triamcinolone	914	0	0.1% (1)	0	Nr			0.7% (1)	0	0.7% (1)	Nr			Nr
BRVO	Parodi 2008 [27]	Triamcinolone	11	0			0			0			Nr			0
				1.25 mg			Grid									
BRVO	Russo 2009 [28]	Bevacizumab	25	0	0		0			Nr			Nr			Nr
BRVO	Donati 2012 [29]	Bevacizumab	63 [*]	Nr			Nr			Nr			Nr			Nr

Nr: Not reported

^{*} Result of calculation as follows: median number of intravitreal injection multiplied by number of participants per group[#] Number of implants by dosing groups: 1342 0.7 mg, 373 0.35 mg

doi: 10.1371/journal.pone.0078538.t003

progression occurred more frequently in the triamcinolone groups than in the observation group. 20% of patients treated with triamcinolone 1 mg and 35% of patients treated with triamcinolone 4 mg received antiglaucomatous medication compared to 8% of patients included in the observation group ($p=0.02$ for observation versus 1 mg, $p<0.001$ for observation versus 4 mg, $p=0.02$ 1 mg versus 4 mg). Glaucoma surgery was performed in 2 patients (2.2%) treated with triamcinolone 1

mg until month 12 and 2 patients (4%) treated with 4 mg between month 12 and 24 [5]. Cataract progression was seen in 26% (1 mg), 33% (4 mg) and 18% (observation) of patients ($p=0.14$, X^2 test). No cataract surgery was performed in the observation group (0%), whereas continuous cataract progression led to surgery in 5.5% of patients treated with triamcinolone 1 mg until month 24. The rate of cataract surgery was considerably higher in the triamcinolone 4 mg group with

Table 4. Ocular safety – IOP, glaucoma, cataract, vitreous hemorrhage.

Study	Investigational product	IOP increase [%], (number of eyes)			Medical treatment of IOP [%], (number of eyes)			Glaucoma surgery [%], (number of eyes)			Cataract progression [%], (number of eyes)			Cataract surgery [%], (number of eyes)			Vitreous hemorrhage [%], (number of eyes)			
		1 mg	4 mg	Obs.	1 mg	4 mg	Obs.	1 mg	4 mg	Obs.	1 mg	4 mg	Obs.	1 mg	4 mg	Obs.	1 mg	4 mg	Obs.	
CRVO	SCORE 2009, 2011 [5,32]	5.4% (5)	8.8% (8)	1.1% (1)	20% (18)	35% (32)	8% (7)	2.2%	0	0	26%	33%	18%	0	4.9% (4)	0	4.3% (4)	0	4.5% (4)	
		12-24 months																		
CRVO/ BRVO	GENEVA 2011 [8,9]	32.8%			25.5%	25%	Nr	1.3% (14)	0		29.8%	19.8%	5.7%	1.3%	1.8%	1.1%	Nr			
		6-12 months																		
CRVO	Epstein 2012 [26]	Nr			Nr			Nr			Nr			Nr			Nr			
		COPERNICUS 2013 [20,36]																		
CRVO	COPERNICUS 2013 [20,36]	12.3%	13.5%		Nr		0	2.7% (2)/ 1.7% (1)		0.9% (1)	1.7% (1)		Nr			0.9% (1)	5.4% (4)/ 1.7% (1)			
		0.3/ 0.5 mg	0.5/ Sham/ 0.5 mg	Sham/ 0.5 mg	0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg	0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg	0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg	0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg	0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg	0.5 mg
CRVO	CRUISE 2011 [22]	Nr			Nr			Nr		3.8% (5)	7% (9)	0/ 1.8% (2)	Nr			5.3% (7)	5.4% (7)	7% (9)/ 1.8% (2)		
		HORIZON 2012 [30]																		
CRVO/ BRVO	HORIZON 2012 [30]	0	0.9% (1)	0	Nr			Nr		5.6% (6)	5.1% (5)	3.1% (3)	Nr			0	0	1% (1)		
		BRAVO 2011 [23]																		
BRVO	BRAVO 2011 [23]	Nr			Nr			Nr		4.5% (6)	6.2% (8)	3.1% (4)/ 2.6% (3)	Nr			5.2% (7)	1.5% (2)	4.6% (1)		
		HORIZON 2012 [30]																		
BRVO/ CRVO	HORIZON 2012 [30]	1% (1)	0.9% (1)	0	Nr			Nr		9.7% (10)	5.8% (6)	6.5% (6)	Nr			1% (1)	0	1.1% (1)		
		1 mg	4 mg	Grid	1 mg	4 mg	Grid	1 mg	4 mg	Grid	1 mg	4 mg	Grid	1 mg	4 mg	Grid	1 mg	4 mg	Grid	Grid
BRVO	SCORE 2009, 2011 [4,32]	Nr			7%	41%	2%	0	0	25%	35%	13%	0	2.9% (4)	2.2% (3)	0.7% (1)	2.2% (3)	1.5% (2)		
		12-24 months																		
BRVO	Parodi 2008 [27]	Nr			54%	0		Nr			Nr			Nr			0			
		SGLT																		
4 mg + SGLT																				

Table 4 (continued).

	Study	Investigational product	IOP increase [%], (number of eyes)	Medical treatment of IOP [%], (number of eyes)	Glaucoma surgery [%], (number of eyes)	Cataract progression [%], (number of eyes)	Cataract surgery [%], (number of eyes)	Vitreous hemorrhage [%], (number of eyes)
BRVO	Russo 2009 [28]	Bevacizumab	0	0	0	0	0	Nr
BRVO	Donati 2012 [29]	Bevacizumab	Nr	Nr	Nr	Nr	Nr	Nr

IOP: intraocular pressure, Nr: not reported, PRN: pro re nata, Obs.: Observation, SGL T: subthreshold grid laser photocoagulation
doi: 10.1371/journal.pone.0078538.t004

4.9% of patients until month 12, and 42% between month 12 and 24 [5].

SCORE-BRVO reported 1 case of endophthalmitis in 914 intravitreal injections (0.1%) [5]. Furthermore, 1 retinal detachment (0.7%) was seen in each of the grid and triamcinolone 1 mg groups. Likewise in CRVO patients, IOP rise (2-41%) and cataract progression (13-35%) were the most frequent ocular adverse events and found most frequently in patients treated with triamcinolone 4 mg. From the patients with increased IOP, 2% (grid), 7% (1 mg) and 41% (4 mg) received antiglaucomatous medication ($p=0.03$ for grid versus 1 mg, $p<0.001$ for grid versus 4 mg and also for 1 mg versus 4 mg, X^2 test), no glaucoma surgery was performed until month 12; however, 2 patients (2.8%) within the triamcinolone 4 mg group underwent surgery until month 24 [4]. Cataract progression occurred in 13% (grid), 25% (1 mg) and 35% (4 mg) of patients ($p=0.03$ for grid versus 1 mg, $p<0.001$ for grid versus 4 mg, $p=0.1$ for 1 mg versus 4 mg) and led to an increasing number of cataract surgeries from year 1 to year 2 in all groups (grid: 2.2% to 7.3%, 1 mg: 0% to 9.5%, and 4 mg: 2.9% to 48.6%) [4].

Considering neovascular events, the SCORE study provided solid and detailed information on year 1 and 2 adverse events [4,5], and looked at 36-month incidences and differences between the treatment groups in an additional report [32]. Neovascularization of the iris (NVI) and neovascular glaucoma (NVG) occurred at a rate of 3.2% (NVI) and 5.8% (NVG) in CRVO (36-month over-all incidence) and at a less frequent rate in BRVO (0.3% NVI versus 2.2% NVG) [32]. The cumulative incidence at 12 and 36 months was 6.1% for NVI and 8.5% for NVG in CRVO, compared to 1.3% NVI and 2.4% NVG in BRVO [32]. The investigators did not find a significant difference between treatment groups (CRVO $p=0.31$, BRVO $p=0.18$). The 36-month incidences for preretinal hemorrhage (PRH) or vitreous hemorrhage (VH) were 7.6% in CRVO and 3.8% in BRVO. The cumulative incidences at 12 and 36 months for neovascularization of the disc or elsewhere (NVD/E) were 2.8% (CRVO) and 2.9% (BRVO) and 8.8% (CRVO) and 7.6% (BRVO) for PRH/E, respectively [32].

Parodi et al. did not see any endophthalmitis in 11 patients treated with a single triamcinolone injection after 12 months [27]. The only ocular adverse event that occurred was a rise in IOP which was treated with antiglaucomatous medication in 54% of the triamcinolone plus SGLT group compared to 0% in the SGLT group [27].

Dexamethasone Implant - The GENEVA study (CRVO/BRVO). Safety outcomes were not stratified into CRVO and BRVO patients but comparison was possible between the 6 different treatment arms (stratified by the number and dose of dexamethasone implant) at month 12 [8]. Haller et al. reported no endophthalmitis and a "rare" occurrence of retinal detachment and retinal tear without giving any numbers. Frequently observed adverse events were IOP rise and cataract progression. The increase of IOP ≥ 10 mmHg was most pronounced at day 60 with 32.8% over all injections (12.6% after the 1st injection and 15.4% after the 2nd, dexamethasone 0.7/0.7 mg group). Taken all dexamethasone treatment groups together, 1.3% of patients underwent

glaucoma surgery compared to 0.0% of patient in the sham/ no implant group. 25.5% (dexamethasone 0.7/0.7 mg), 25.0% (dexamethasone 0.35/0.7 mg), and 28.1% (sham/ dexamethasone 0.7 mg) received medical treatment for IOP compared to 0% in 0.35 mg/no implant group and sham/no implant group [8]. Cataract progression varied between 5.7% (sham/no implant), 10.5% (sham/0.7 mg) and 29.8% (0.7/0.7 mg) dexamethasone treated patients. There was no safety information on vitreous hemorrhage or neovascularization [8].

Ranibizumab - CRUISE (CRVO), BRAVO (BRVO), and HORIZON trials (CRVO/ BRVO). The incidence of endophthalmitis was 0% in all groups of the CRUISE and BRAVO trials during year 1 [22,23]. During year 2, two patients (0.9%) of the ranibizumab (0.3/0.5 mg) group suffered from endophthalmitis [30]. Retinal tears and retinal detachment were rare among all groups (rate ranged between 0.7% and 1.8%). IOP increase was not reported during the first year [22,23] and rare during the second year [30] (0.9% of each group [CRUISE 0.5/0.5 mg, BRAVO 0.3/0.5 mg and 0.5/0.5 mg] showed an increased IOP). No surgeries and no medical treatments of IOP were reported. No cataract surgery was performed and progression of cataract was low. CRUISE found 3.8% (0.3/0.5 mg ranibizumab) and 7.0% (0.5/0.5 mg) cataract progression within 1 year compared to 1.8% (sham/0.5 mg ranibizumab) [22]. BRAVO found comparable low rates of 4.5% (ranibizumab 0.3/0.5 mg), 6.2% (0.5/0.5 mg) and 5.7% (sham/0.5 mg) [23]. During the second year rates remained fairly stable [30].

Bevacizumab - Epstein 2012 (CRVO), Russo 2009 and Donati 2012 (BRVO). Epstein et al. did not report any endophthalmitis, nor retinal tear or detachment in both groups [26]. In the sham group 16.7% developed NVI, but no one in the bevacizumab group. By month 12 no new development of NVI was found in either group. Furthermore, in all 16.7% NVI cases, neovascularization regressed completely until month 12 after treatment changed from sham to bevacizumab 1.25 mg [26].

Russo et al. and Donati et al. reported no ocular adverse events in 25 intravitreal injections of bevacizumab [28,29]. Neither group showed complications during grid laser photocoagulation [28].

Aflibercept - The COPERNICUS trial (CRVO). One of 189 patients presented with infectious endophthalmitis after intravitreal aflibercept 2 mg (0.9% [aflibercept], and 0% [sham/ aflibercept]) [20]. Retinal tear was rare and occurred in 1 patient (1.7%) of the sham/aflibercept group. Vitreous hemorrhage was seen in 0.9% (aflibercept) compared to 7.1% (sham/aflibercept, 5.4% at 6 months and 1.7% until 12 months). NVI was found in 2.7% and NVG in 6.8% of control patients, whereas none of the aflibercept 2 mg group showed neovascular changes. Three patients (4.4%) of the sham/ aflibercept group (2.7% at 6 months and 1.7% at 12 months) underwent glaucoma surgery, but none of the aflibercept patients. Increased IOP was equally distributed between both groups (12.3% [aflibercept] and 13.5% [sham/aflibercept]), as well as cataract progression (0.9% and 1.7%, respectively) [20].

Table 5. Systemic safety.

Study		Investigational product	Death (any cause) [%], (number)			Myocardial infarction [%], (number)			Cerebrovascular accident [%], (number)			Nonocular hemorrhage [%], (number)			Infections [%], (number)		
			1 mg	4 mg	Obs.	1 mg	4 mg	Obs.	1 mg	4 mg	Obs.	1 mg	4 mg	Obs.	1 mg	4 mg	Obs.
CRVO	SCORE 2009, 2011 [5,32]	Triamcinolone	1.1% (1)	2.2% (2)	0	Nr			Nr			Nr			15%	19%	10%
	12-24 months		1% (1)	3.4% (3)	1.1% (1)												
CRVO/ BRVO	GENEVA 2011 [8]	Dexamethasone implant	Nr			Nr			Nr			Nr			Nr		
									1.25 mg	Sham/ 1.25 mg							
CRVO	Epstein 2012 [26]	Bevacizumab	Nr			Nr			0		3.3% (1)	Nr			Nr		
			2 mg	Sham/ 2 mg		2 mg	Sham/ 2 mg		2 mg	Sham/ 2 mg		2 mg	Sham/ 2 mg		2 mg	Sham/ 2 mg	
CRVO	COPERNICUS 2013 [20]	Aflibercept	0	2.7% (2)		0.9% (1)	1.7% (1)		0	1.7% (1)		Nr			7.9% NPG, 7.9% URI	6.8% NPG, 5.4% URI	
			0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg	0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg	0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg	0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg	0.3/ 0.5 mg	0.5/ 0.5 mg	Sham/ 0.5 mg
CRVO	CRUISE 2011 [22]	Ranibizumab	0	0.8% (1)	0	0.8% (1)	0.8% (1)	0.8% (1)	0.8% (1)	1.6% #	0	0			Nr		
CRVO/ BRVO	HORIZON 2012 [30]	Ranibizumab	1% (1)	3% (3)	3.1% (3)	0.9% (1)	0	0	0.9% (1)	2.1% (2)	1% (1)	0	2% (2)	2.1% (2)	7.3% (7) NPG	7.1% (7) NPG	7.3% (7) NPG
BRVO	BRAVO 2011 [23]	Ranibizumab	0	0	0	0	0.8% (1)	0.9% (1)	0.7% (1)	0.8% (1)	0.8% (1)	1.5% (2)	0.8% (1)	0	Nr		
BRVO/ CRVO	HORIZON 2012 [30]	Ranibizumab	1% (1)	2.9% (3)	0	1% (1)	1% (1)	0	2.9% (1)	0	1% (1)	0	2.9% (3)	0	5.3% NPG	1% (1) NPG	7.5% (7) NPG
			1 mg	4 mg	Grid	1 mg	4 mg	Grid	1 mg	4 mg	Grid	1 mg	4 mg	Grid	1 mg	4 mg	Grid
BRVO	SCORE 2009, 2011 [4,32]	Triamcinolone	2.2% (3)	1.4% (2)	2.1% (3)	Nr			Nr			Nr			16%	15%	10%
			(2)	(7)	(2)												
BRVO	Parodi 2008 [27]	Triamcinolone	Nr			Nr			Nr			Nr			Nr		
BRVO	Russo 2009 [28]	Bevacizumab	Nr			Nr			Nr			Nr			Nr		
BRVO	Donati 2012 [29]	Bevacizumab	Nr			Nr			Nr			Nr			Nr		

Obs.: Observation, Nr: not reported, NPG: Nasopharyngitis, URI: upper respiratory infection

sum of one TIA (0.8%) and one ischemic stroke (0.8%)

doi: 10.1371/journal.pone.0078538.t005

Study results: Safety Outcome – systemic adverse events (Table 5)

In general, the incidence of systemic adverse events was low and did not significantly differ between treatment groups in the multicenter trials [4,5,8,20,22,23,30]. The single center trials did not report any non-ocular adverse events [27–29].

Triamcinolone - The SCORE study (CRVO/ BRVO). In the SCORE-CRVO study 1.1% of patients (1 mg), 2.2% (4 mg) and none (observation) died within 12 months. During the second

year rates of deaths remained low with 1% (1 mg), 3.4% (4 mg), and 1.1% (observation) [5]. Death originated from different causes: among cardiovascular incidences, 1 myocardial infarction (4 mg) and 1 brain hemorrhage (1 mg) led to death. Infections were equally distributed between groups (15% [1 mg], 19% [4 mg] and 10% (observation), respectively [5]. In comparison, SCORE-BRVO reported 7 deaths equally distributed between groups (Table 5) [4]. The class of infections was found most frequently among systemic adverse

events, i.e. 10% (grid), 16% (1 mg) and 15% (4 mg; not significant).

Dexamethasone Implant - The GENEVA study (CRVO/ BRVO). No systemic adverse events are reported in this trial [8].

Ranibizumab - CRUISE (CRVO), BRAVO (BRVO) and HORIZON trial (CRVO/ BRVO). CRUISE investigators found an incidence of non-ocular serious adverse events potentially related to VEGF inhibition of 1.6% (in patients treated with ranibizumab 0.3/0.5 mg), 3.2% (in patients treated with 0.5/0.5 mg) and 1.6% (of sham treated patients) until month 6. Systemic adverse events were not reported in the sham/0.5 mg PRN group until month 12 [22]. Death by vascular causes did not occur among all groups and death from unknown cause was found in 0.8% in the ranibizumab 0.5/0.5 mg group. One patient of each group suffered from myocardial infarction (0.8%). Cerebrovascular events were rare with 0.8% (0.3/0.5 mg), 1.6% (0.5/0.5mg), and 0% (sham/0.5 mg) [22]. Non-ocular hemorrhages did not occur and infections were not reported [22].

BRAVO investigators reported an incidence of non-ocular serious adverse events potentially related to VEGF inhibition of 4.5% (ranibizumab 0.3/0.5 mg), 4.6% (0.5/0.5 mg), 0.8% (sham until month 6) and 1.7% (sham/0.5 mg) until month 12 [23]. Myocardial infarction occurred in 0% (0.3/0.5 mg), 0.8% (0.5/0.5 mg) and 0.9% (sham/0.5 mg). One of each group suffered from stroke (0.7–0.8%). In 1.5% (0.3/0.5 mg), 0.8% (0.5/0.5 mg), and 0% (sham/0.5 mg) of patients non-ocular hemorrhage was reported [23]. Rates remained low during the course of year 2 (0–3%) [30]. Rates of infections, i.e. nasopharyngitis, are given for HORIZON patients and were 7% in every CRVO group, but differed among BRVO groups between 1% (0.5/0.5 mg), 5.3% (0.3/0.5 mg), and 7.5% (sham/0.5 mg) [30].

Bevacizumab - Epstein (CRVO). Epstein et al. did not report serious non-ocular adverse effects [26,31]. However, one patient (3.3%) from the sham/bevacizumab group suffered from a transient ischemic attack and dropped out of the study.

Aflibercept - The COPERNICUS trial (CRVO). During the first 12 months of the ongoing COPERNICUS trial, no deaths were reported in the aflibercept 2 mg group. However, 2 deaths (2.7%) occurred in the sham/aflibercept group [20]. Both deaths had a vascular cause (myocardial infarction and cardiac arrhythmia). Non-lethal myocardial infarction was seen in 1 patient of each group (0.9% and 1.7%). The most frequent infections were nasopharyngitis (7.9% [aflibercept] and 6.8% [sham/aflibercept]) and upper respiratory infections (7.9% versus 5.4%). Overall non-ocular serious adverse events were rare and similarly distributed between groups (5.3% [aflibercept] versus 8.1% [sham] at six months, and 6.4% [aflibercept] versus 8.3% [sham/aflibercept] at 12 months) [20]. The incidence of systemic adverse events did not differ significantly between groups.

Methodological Quality and Risk of Bias (Table 6)

Adequate blinding over 12 months was reported in 3 trials (COPERNICUS, CRUISE, BRAVO) [20,22,23]. SCORE investigators were blind to the dosage (triamcinolone 1 mg

versus 4 mg), but not to observation [4,5]. Two studies changed the status of blinding after the primary end point was reached from double-blind to open-label at month 6 (GENEVA, Epstein) [8,26]. HORIZON followed patients from CRUISE and BRAVO trials in an open-label design after month 12 [30]. However, interpretation of data is limited by the high drop-out rate in all groups and the early termination after approval for ranibizumab. Validity of efficacy was high in 6 trials based on adequate blinding, transparency of patient flow, comparability between groups, sample size calculation and ITT-analysis (SCORE-BRVO/ -CRVO, Epstein, COPERNICUS, CRUISE and BRAVO).

Based on our quality assessment, validity of safety was high in 4 trials (SCORE-BRVO/ -CRVO, GENEVA, and COPERNICUS) [4,5,8,20]. Four other trials showed moderate validity due to incomplete definition of expected adverse events or methods used to collect data on adverse events (Epstein, CRUISE, BRAVO, HORIZON) [22,23,26,30]. Information on ocular and systemic adverse events varied between very detailed reporting within most multicenter trials (including additional information available as online supplements) and, on the other hand, the overall notification of “no ocular and systemic adverse events” in a single-center trial.

The trials by Parodi et al., Russo et al. and Donati et al. substantially lack quality (unclear or no blinding, moderate (low) transparency of patient flow, lack of information on sample size calculation, data analysis, adverse events collection and reporting) [27–29]. Additionally, comparability of groups is impaired (due to small sample size) [27–29].

Discussion

Principal Findings

Visual acuity. All anti-VEGF agents show a better gain of BCVA than steroids at month 12 (Figure 2 CRVO, 3 BRVO). A significant improvement of BCVA after 12 months of intravitreal anti-edematous therapy in CRVO is achieved with aflibercept 2 mg every 4 weeks for 6 months followed by a monthly PRN scheme (+16.2 letters, 8.5 injections) as well as continuous intravitreal injections of bevacizumab 1.25 mg every 6 weeks (+16.1 letters, 8 injections) [20,26]. In addition, ranibizumab (0.5 mg) leads to an improvement of VA of +13.9 letters in fixed plus monthly PRN dosage regimen as shown in the CRUISE study [22].

The SCORE-CRVO and GENEVA studies resulted in a stabilization of BCVA at month 12 with considerable less frequent injections (SCORE: -1.2 letters [2 injections], GENEVA: approximately +2 letters [1.8 injections]) [5,8]. Improvement of BCVA in triamcinolone treated patients was significantly better than in the observation group [5].

In BRVO best improvement of BCVA was found for ranibizumab in BRAVO at month 12 [23]. Treatment with 0.5 mg ranibizumab monthly for 6 months followed by monthly PRN resulted in +18.3 letters (8.4 injections) [23]. In comparison, bevacizumab seems to achieve similar results (+15 to 17 letters), but interpretation is limited by the very low numbers of participants in the included RCTs (9 to 15 patients per group) [27–29]. The RCT for aflibercept in BRVO is not yet

Table 6. Methodological Quality and Risk of Bias.

	Study	adequate blinding	transparency of patient flow	Comparability		ITT-analysis	Validity efficacy**	definition of definition of method used expected to collect AE Validity safety***		
				between groups	sample size calculation			AE	data	***
CRVO	SCORE 2009, 2011 [5,32]	No *	Yes	Yes	Yes	Yes	High	Yes	Yes	High
BRVO	SCORE 2009, 2011 [4,32]	No *	Yes	Yes	Yes	Yes	High	Yes	Yes	High
CRVO/BRVO	GENEVA 2011 [8,9]	No ##	Yes	Yes	No	unclear +	Moderate - low	Yes	Yes	High
CRVO	Epstein 2012 [26,31]	No ##	Yes	Yes	Yes	Yes	High	No	Yes	Moderate
CRVO	COPERNICUS 2013 [20,36]	Yes	Yes	Yes	Yes	Yes	High	Yes	Yes	High
CRVO	CRUISE 2011 [22]	Yes	Yes	Yes	Yes \$	Yes	High	No	Yes	Moderate
BRVO	BRAVO 2011 [23]	Yes	Yes	Yes	Yes \$	Yes	High	No	No	moderate
CRVO/BRVO	HORIZON 2012 [30]	No	Yes	Yes	No	Yes	Moderate	No	Yes	Moderate
BRVO	Parodi 2008 [27]	Unclear §	Yes ++	Yes	No	Unclear	Low	No	No	Low
BRVO	Russo 2009 [28]	No	Yes ++	Yes	No	Unclear	Low	No	No	Low
BRVO	Donati 2012 [29] #	No	Yes ++	Yes	No	Unclear	Low	No	No	Low

ITT: Intention-to-treat, AE: adverse event

Studies with investigational product in treatment and control group (i.e. control = combination therapy)

** Validity of efficacy is based on: adequate blinding, transparency of patient flow, comparability between groups, sample size calculation and ITT-analysis

*** Validity of safety is based on: definition of expected AE, definition of method used to collect AE data, blinding, transparency of patient flow, and comparability between groups

* Blind to dosage of triamcinolone (1 mg or 4mg), but not to treatment arms (observation versus triamcinolone)

Double-masked trial for 6 months, followed by open-label for 6 months

§ patient blind regarding intravitreal injection (ranibizumab or sham), unclear, if treating and examiners blind

\$ Study was not powered to compare efficacy outcomes at 6 months, efficacy analysis based on descriptive statistics, 12 months data post hoc analysis

+ Results of the 2 trials were pooled, results in the open label extension were analyzed for all patients according to the actual treatment received

++ description of all patients all visits within results, but no comment within statistics

doi: 10.1371/journal.pone.0078538.t006

published ("Study to Assess the Clinical Efficacy and Safety of VEGF Trap-Eye (Intravitreal Aflibercept Injection [IAI]), Also Commercially Known as EYLEA™ in Patients With Branch Retinal Vein Occlusion (BRVO)", NCT01521559, Regeneron Pharmaceuticals, source: clinicaltrials.gov). A multitude of non-RCTs is published on bevacizumab in RVO, but lack of quality, especially in adverse events documentation and reporting, low numbers of participants and multitude of therapeutic strategies lead to exclusion within this review.

BCVA results at 12 months were less favorable for steroids in BRVO. No significant effect was found for triamcinolone in SCORE-BRVO versus grid (2.1 injections) [4]. BRVO patients treated with dexamethasone implant deteriorated to approximately +6 letters after temporary visual improvement at month 12 (1.8 injections) [8].

Grid laser photocoagulation. Despite the gold standard of grid laser photocoagulation implemented after the BVOS [2], we found heterogeneous control groups ranging from sham injection in BRAVO to grid laser photocoagulation in SCORE-BRVO [4,23]. BRAVO allowed grid as rescue therapy, which complicates the indirect comparison of the trials. Only one trial compared grid laser photocoagulation in combination to

intravitreal bevacizumab 1.25 mg to bevacizumab alone, and found even better results in the combination group (+20 letters [median 3 injections]) [29]. There are currently 4 trials registered that compare intravitreal ranibizumab with macular laser photocoagulation and/ or a combination therapy (RABAMES, NCT00562406; BRIGHTER, NCT01599650; NCT01189526, clinicaltrials.gov, and EUCTR2008-007175-24-HU, www.who.int/ictpr/en/), and 1 RCT comparing triamcinolone versus grid versus combination (ChiCTR-TRC-05000661, www.who.int/ictpr/en/). A major limitation of all studies with grid laser photocoagulation is the heterogeneity of the treatment itself and the lack of evaluation of grid by the reading centers.

Two year data. Two year data are only published for ranibizumab within the course of the HORIZON trial [30]. But interpretation of data is limited by the high drop-out rate in all groups and the early termination after approval for ranibizumab. Patients received significantly less intravitreal injections during the second year (CRVO 3.5 injections, BRVO 2.1) and lost little visual acuity BCVA (CRVO -4 letters, BRVO -0.7 letters) [30]. Two years follow-up of COPERNICUS [20]

and GALILEO [33] will provide us with more evidence regarding anti-VEGF in CRVO.

Clinically significant gain in BCVA (more than 15 letters). Analysing SCORE, COPERNICUS, CRUISE, BRAVO and Epstein studies, anti-VEGF treatment of macular edema leads to a clinically significant visual improvement in almost twice as many participants as triamcinolone or grid laser photocoagulation. Approximately 25% responders were found for triamcinolone [4,5], while anti-VEGF trials found 50% [22], 55% [20] and 60% [26] in CRVO (Figure 2) and 60% in BRVO [23] (Figure 3). Results at 12 months in control groups are better than after 6 months and the differences between drug and control groups smaller, because patients were treated open-label with the investigational product for months 6 to 12 in all multicenter trials (CRUISE, BRAVO, Epstein, COPERNICUS, GENEVA) [8,20,22,23,26]. Our analysis of 12 months data still shows a striking difference over all anti-VEGF trials between the groups, which were treated with the individual therapeutic agent from the beginning and the control/PRN groups, which had a delay of 6 months before therapy started. In conclusion, the effect of the anti-VEGF agents on macular edema in RVO is more pronounced if treatment starts early after onset of macular edema.

Time between symptoms and treatment. On basis of these results, one can assume that time between occlusion and treatment is a critical factor both for patient choice in the context of clinical trials but also with regard to the therapeutic effect and the interpretation of study data. In practice, studies including and treating patients shortly after RVO onset might find better visual acuity outcome and/ or response to treatment than trials including old RVO with persistent macular edema, resulting in false negatively reduced outcome. Therefore, we specifically depicted time of occlusion in the list of study characteristics. Both, mean time from occlusion and percentage of patients with ≤ 2 -3 months occlusion time differed between the included RCTs. One explanation is the choice of time span for inclusion criteria. In CRVO trials, mean time of occlusion ranged from 1.9 months (bevacizumab, Epstein 2012), 1.9 and 2.7 (aflibercept, COPERNICUS), 3.3 (ranibizumab, CRUISE), to 4.2 (triamcinolone, SCORE) and 5.2 months (dexamethasone implant, GENEVA) (Table 1) [4,5,8,20,22,26]. We found that the time to treatment is considerably longer in the trials investigating steroids compared to anti-VEGF trials. This is also expressed in high percentages of "fresh" RVO ≤ 2 months of 56-71% in COPERNICUS, RVO ≤ 3 months of 69% in CRUISE and only 15-18% in GENEVA (Table 1) [8,20,22]. The same discrepancy was found in BRVO trials, the percentage of "fresh" RVO was 65% in BRAVO and 37% in SCORE [23,34]. This might put into perspective the differences in visual acuity outcome between the RCTs.

Ocular adverse events. The rate of endophthalmitis was very low in all RCTs ($<1\%$) regardless of the substance class. SCORE-BRAVO, GENEVA, Epstein, and CRUISE did not find any endophthalmitis in approximately 6000 intravitreal injections (Table 3) [4,8,22,26]. These findings in RVO are in accordance to safety results and analysis of ranibizumab and bevacizumab in AMD [35]. Due to the different frequency of

injection between anti-VEGF (8 injections/ 12 months) and steroids (2 injections/ 12 months) the risk of endophthalmitis for the individual patient is higher in an anti-VEGF treatment regimen than in triamcinolone or dexamethasone.

IOP increase and medical treatment of IOP occurred more frequently in SCORE and GENEVA than in any anti-VEGF trial (Table 4). Looking into SCORE data, IOP increase seems to be dose-dependent with maximum rates of 35% (CRVO) and 41% (BRVO) for triamcinolone 4 mg [4,5]. GENEVA found approximately 25% in all dexamethasone implants after 6 months. It remains unclear, whether the 10% rate reported for 6-12 months data are in addition to these 25% [8,9]. No need for medical or surgical IOP control was reported in Epstein, CRUISE, BRAVO and HORIZON (Table 4) [22,23,26,30]. COPERNICUS found IOP increase in all groups of $> 10\%$, which needed no medical treatment [20,36]. Rates of glaucoma surgery was occasional in GENEVA (1.3%, all dexamethasone implants), low in SCORE (2.2% CRVO triamcinolone 4 mg) and COPERNICUS (2.7% sham group, 1.7% sham/aflibercept 2 mg), and not reported for bevacizumab or ranibizumab trials [5,8,20,22,23,26]. It is remarkable, that in SCORE and GENEVA trials only participants with treatment of steroids developed glaucoma which needed surgery, while in anti-VEGF trials participants did not need surgery under ranibizumab or bevacizumab treatment [4,5,8,22,23,26]. Additionally, in COPERNICUS, surgery had to be performed in the sham/aflibercept group, probably due to secondary neovascular glaucoma [20].

Rates of cataract and surgery were found more frequently in SCORE and GENEVA with 20 - 35% cataract progression, compared to none in anti-VEGF trials (Table 4) [4,5,8].

Vitreous hemorrhages, either due to the injection procedure or because of progression to neovascular RVO, was found seldom in all treatment groups. We found a tendency towards higher rates in control groups in CRVO, which received treatment 6 months later. This difference was most pronounced in COPERNICUS (0.9% [aflibercept] versus 7.1% [sham/aflibercept]), and less striking in CRUISE (5.3% [ranibizumab 0.3/0.5 mg], 5.4% [0.5/0.5 mg] versus 8.8% [sham/0.5 mg]) [20,22]. GENEVA did not report vitreous hemorrhages [8,9]. In BRVO no difference between treatment and control groups was found [23,30]. This finding may be explained by the higher risk of CRVO for vitreous hemorrhage compared to BRVO due to the bigger area of tissue affected [1,2].

In conclusion, physicians have to take these differences in ocular risk profiles between anti-VEGF and steroids into account to choose the best therapy for their individual patients.

Systemic adverse events. The systemic risk seems to be comparable low between all compounds given intravitreally in RVO. Rates of death were low (0-3%) over all groups and no statistically significant difference was found (Table 5). Comparison of systemic adverse event rates was biased by different reporting and difficulty in data extraction. None of the RCTs investigated in this review found significant differences within cerebrovascular events, non-ocular hemorrhages or infections. Infections were found in up to 7.9% in anti-VEGF trials (all groups, all substances), but more often in SCORE patients (10-19% all groups) [4,5,20,30]. Rates of infections are

higher than reported before in ranibizumab (2-3.8%) or bevacizumab (4-6%) [14,35]. One has to take into account that RVO patients tend to be younger than AMD patients, but present more often with cardiovascular diseases [37]. Hypertension, diabetes, hypercholesterolemia and adiposity contribute to RVO but not to AMD [37,38]. Thus it is almost surprising that RVO trials do not find more myocardial infarction or cerebrovascular events than AMD trials. A possible explanation might be age as risk factor for both. Another explanation is the selection bias at time of inclusion within a RCT. Individuals presenting with considerable systemic diseases in their recent past medical history are most often excluded. Certainly, in all trials numbers of participants are still too low to be calculated to detect small differences in rare (systemic) events.

Strengths and Limitations. We limited our search to RCTs and outcome data of at least 12 months to strengthen the implication for clinical practice as RVO treatment seeks an ongoing intervention. To investigate efficacy and safety outcome on best available data, we just referred to RCTs. We anticipated a vast number of reports of off-label bevacizumab use in RVO similar to the publication mode of AMD trials. Regarding safety data, previous analysis of similar reports often found a lack of monitoring and reporting of adverse events [35]. These shortcomings were also found in our review. Overall, we found better reporting and analysis in multicenter RCTs than in monocenter trials.

Our search detected no head-to-head studies comparing two different anti-VEGF agents or anti-VEGF vs. steroids. This is a major limitation when comparing the therapeutic outcome and prevents us from taking clear recommendation for clinical choice. Only direct comparison within the same trial leads to sufficient evidence due to the same criteria for all participants. Regarding RVO, comparison of baseline characteristics is crucial due to the pathophysiology and natural course of macular edema in RVO. In our opinion, time between occlusion and start of treatment is a major confounder comparing published trials. As discussed above, the differences in rates of “young” versus “old” RVO between anti-VEGF and steroids treated patients may contribute to the differences seen in outcome. Data extraction was biased by different or no reporting of important outcomes, like the percentage of patients with gain or loss of 15 letters at 12 months. This is partly due to the different choice of primary end points between the trials. Complicating data retrieval, some data on CRVO and BRVO were published combined, e.g. outcome of dexamethasone treatment [8]. These limitations can only be overcome by trials comparing 2 or more substances directly. Fortunately, several head-to-head trials started recently to investigate dexamethasone implant versus bevacizumab (NCT01231633, www.who.int/ictcp/en/) or ranibizumab (COMRADE-B/-C/Extension, NCT01396057, NCT01396083, NCT01580020, and COMO, NCT01427751, www.who.int/ictcp/en/). Notably, COMRADE trials investigate BRVO and CRVO separately, and COMO is a head-to-head trial for BRVO. Additionally, we found two trials addressing the question of costs and effectiveness comparing bevacizumab vs. ranibizumab in RVO (NCT01428388, NCT01635803, www.who.int/ictcp/en/).

On the other hand, we see the strengths of our systematic review in the critical and detailed data extraction of all 12 months data on anti-VEGF and steroids available in the current literature. Analysis of the detected RCTs shows both remarkably high validity of efficacy and safety for many of these trials which strengthens the quality of reported outcomes and supports comparison of outcomes. To our knowledge this review is the first to report and compare 12 months safety and efficacy data of five therapeutic agents currently in use for macular edema in RVO. This is of great value to all ophthalmologists to help decide between the therapeutic options and to calculate the accompanying risks.

Other reviews. We identified only 1 Cochrane review on anti-VEGF for macular edema secondary to BRVO [39] and 1 review evaluating CRVO [40]. Both reviews limited their search and outcome reporting on anti-VEGF agents without considering steroids. In addition, their results are based on data of only 2 (quasi-)RCTs. In contrast, our systematic review includes 11 up-to-date RCTs investigating and comparing different therapeutic (anti-VEGF and steroids) options with follow-ups of at least 12 months. Therefore, the current systematic review functions as a detailed update on knowledge of treatment for macular edema in both, BRVO and CRVO.

Implications for Clinical Practice. Our search detected 5 trials in CRVO, each of them providing evidence for the superiority of triamcinolone, dexamethasone implant, bevacizumab, ranibizumab, and aflibercept, respectively [5,8,20,22,26], in comparison to observation or sham, which was the unsatisfying gold standard before initiating intravitreal therapy [1]. In BRVO, ranibizumab has been established as the new gold standard, even if it received FDA approval without testing it against the grid laser photocoagulation (the previous gold standard for this indication) [2,23]. Smaller trials implicate superiority of bevacizumab [25,26] or ranibizumab (RABAMES, www.clinicaltrials.gov) against grid [28,29]. But more evidence is needed to judge, whether patients may profit from a combination or change in therapeutic strategy.

Regarding head-to-head studies, we look forward to ongoing clinical trials such as COMO and COMRADE, comparing ranibizumab directly with dexamethasone implant (www.clinicaltrials.gov). We also expect vast use of off-label bevacizumab due to monetary necessities as seen in AMD and appreciate scientific evidence such as given by Epstein for bevacizumab in CRVO [26,35,41,42]. Given the different medical background of RVO compared to AMD patients, we should not uncritically transfer results from comparative CATT and IVAN trials on the question of off-label use of bevacizumab versus ranibizumab in RVO [14,15]. RVO patients may present with a different spectrum of underlying diseases and potentially higher risk profiles. Evidence from head-to-head studies addressing the question of costs and effectiveness comparing bevacizumab versus ranibizumab in RVO (NCT01428388, NCT01635803, www.who.int/ictcp/en/) is needed to judge upon safety and efficacy in RVO.

In conclusion, macular edema in CRVO and BRVO responds to intravitreal therapy of steroids and various anti-VEGF agents. Best visual acuity results at 1 year are found after aflibercept 2 mg and bevacizumab 1.25 mg in CRVO, and

ranibizumab 0.5 mg in BRVO. Cataract and glaucoma are the main draw-backs in the use of triamcinolone and dexamethasone, while low injection frequency is in favor of steroids. To our knowledge, this review is the first to report 12 months data of five different therapeutic intravitreal agents currently used in RVO. Therefore, these findings may be of great value for all ophthalmologists.

Supporting Information

Text S1. Search strategy in Medline (Ovid).
(RTF)

Table S1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist.

Literature

1. Natural history and clinical management of central retinal vein occlusion. The Central Vein Occlusion Study Group (1997). *Arch Ophthalmol* 115: 486–491. doi:10.1001/archophth.1997.01100150488006. PubMed: 9109757.
2. the Branch Vein Occlusion Study Group (1984) Argon laser photocoagulation for macular edema in branch vein occlusion.. *Am J Ophthalmol* 98: 271–282. PubMed: 6383055.
3. Jonas JB, Kreissig I, Degenring RF (2002) Intravitreal triamcinolone acetate as treatment of macular edema in central retinal vein occlusion. *Graefes Arch Clin Exp Ophthalmol* 240: 782–783. doi: 10.1007/s00417-002-0529-0. PubMed: 12271378.
4. Scott IU, Ip MS, VanVeldhuisen PC, Oden NL, Blodi BA et al.; (2009) A randomized trial comparing the efficacy and safety of intravitreal triamcinolone with standard care to treat vision loss associated with macular edema secondary to branch retinal vein occlusion: the Standard Care vs Corticosteroid for Retinal Vein Occlusion (SCORE) Study report 6 *Arch Ophthalmol* 127: 1115–1128 doi:10.1001/archophth.2009.233.
5. Ip MS, Scott IU, VanVeldhuisen PC, Oden NL, Blodi BA et al.; (2009) A randomized trial comparing the efficacy and safety of intravitreal triamcinolone with observation to treat vision loss associated with macular edema secondary to central retinal vein occlusion: the Standard Care vs Corticosteroid for Retinal Vein Occlusion (SCORE) Study report 5 *Arch Ophthalmol* 127: 1101–1114 doi:10.1001/archophth.2009.234.
6. Jain N, Stinnett SS, Jaffe GJ (2012) Prospective study of a fluocinolone acetate implant for chronic macular edema from central retinal vein occlusion: thirty-six-month results. *Ophthalmology* 119: 132–137. doi: 10.1016/j.ophtha.2011.06.019. PubMed: 21924503.
7. Kube T, Sutter M, Trittel R, Feltgen N, Hansen LL et al. (2006) Carboxymethylcellulose as a new carrier substance for intravitreal injection of reproducible amounts of triamcinolone. *Graefes Arch Clin Exp Ophthalmol* 244: 1385–1390. doi:10.1007/s00417-006-0326-2. PubMed: 16715253.
8. Haller JA, Bandello F, Belfort R Jr, Blumenkranz MS, Gillies M et al. (2011) Dexamethasone intravitreal implant in patients with macular edema related to branch or central retinal vein occlusion twelve-month study results. *Ophthalmology* 118: 2453–2460. doi:10.1016/j.ophtha.2011.05.014. PubMed: 21764136.
9. Haller JA, Bandello F, Belfort R Jr, Blumenkranz MS, Gillies M et al. (2010) Randomized, sham-controlled trial of dexamethasone intravitreal implant in patients with macular edema due to retinal vein occlusion. *Ophthalmology* 117: 1134–1146.e3 doi:10.1016/j.ophtha.2010.03.032. PubMed: 20417567.
10. Ng EWM, Shima DT, Calias P, Cunningham ET Jr, Guyer DR et al. (2006) Pegaptanib, a targeted anti-VEGF aptamer for ocular vascular disease. *Nat Rev Drug Discov* 5: 123–132. doi:10.1038/nrd1955. PubMed: 16518379.
11. Chakravarthy U, Adamis AP, Cunningham ET Jr, Goldbaum M, Guyer DR et al. (2006) Year 2 efficacy results of 2 randomized controlled clinical trials of pegaptanib for neovascular age-related macular degeneration. *Ophthalmology* 113: 1508: e1–25. doi:10.1016/j.ophtha.2006.02.064. PubMed: 16828500.
12. Wroblewski JJ, Wells JA 3rd, Adamis AP, Buggage RR, Cunningham ET Jr et al. (2009) Pegaptanib sodium for macular edema secondary to central retinal vein occlusion. *Arch Ophthalmol* 127: 374–380. doi: 10.1001/archophth.2009.14. PubMed: 19365011.
13. Wroblewski JJ, Wells JA 3rd, Gonzales CR (2010) Pegaptanib sodium for macular edema secondary to branch retinal vein occlusion. *Am J Ophthalmol* 149: 147–154. doi:10.1016/j.ajo.2009.08.005. PubMed: 19875087.
14. Martin DF, Maguire MG, Fine SL, Ying G, Jaffe GJ et al. (2012) Ranibizumab and bevacizumab for treatment of neovascular age-related macular degeneration: two-year results. *Ophthalmology* 119: 1388–1398. doi:10.1016/j.ophtha.2012.03.053. PubMed: 22555112.
15. Chakravarthy U, Harding SP, Rogers CA, Downes SM, Lotery AJ et al. (2012) Ranibizumab versus bevacizumab to treat neovascular age-related macular degeneration: one-year findings from the IVAN randomized trial. *Ophthalmology* 119: 1399–1411. doi:10.1016/j.ophtha.2012.04.015. PubMed: 22578446.
16. Jaisale GB, Szurman P, Feltgen N, Spitzer B, Pielen A et al. (2011) Predictive factors for functional improvement after intravitreal bevacizumab therapy for macular edema due to branch retinal vein occlusion. *Graefes Arch Clin Exp Ophthalmol* 249: 183–192. doi: 10.1007/s00417-010-1470-2. PubMed: 21337042.
17. Stahl A, Struheim I, Hansen LL, Agostini HT, Feltgen N (2010) Bevacizumab in central retinal vein occlusion: a retrospective analysis after 2 years of treatment. *Eur J Ophthalmol* 20: 180–185. PubMed: 19927266.
18. Kondo M, Kondo N, Ito Y, Kachi S, Kikuchi M et al. (2009) Intravitreal injection of bevacizumab for macular edema secondary to branch retinal vein occlusion: results after 12 months and multiple regression analysis. *Retina (Phila, Pa)* 29: 1242–1248. doi:10.1097/IAE.0b013e3181aa8e20. PubMed: 19672216.
19. Holash J, Davis S, Papadopoulos N, Croll SD, Ho L et al. (2002) VEGF-Trap: a VEGF blocker with potent antitumor effects. *Proc Natl Acad Sci U S A* 99: 11393–11398. doi:10.1073/pnas.172398299. PubMed: 12177445.
20. Brown DM, Heier JS, Clark WL, Boyer DS, Vittit R et al. (2013) Intravitreal Aflibercept Injection for Macular Edema Secondary to Central Retinal Vein Occlusion: 1-Year Results From the Phase 3 COPENNICUS Study. *Am J Ophthalmol* 155: 429–437.e7 doi:10.1016/j.ajo.2012.09.026. PubMed: 23218699.
21. Papadopoulos N, Martin J, Ruan Q, Rafique A, Rosconi MP et al. (2012) Binding and neutralization of vascular endothelial growth factor (VEGF) and related ligands by VEGF Trap, ranibizumab and bevacizumab. *Angiogenesis* 15: 171–185. doi:10.1007/s10456-011-9249-6. PubMed: 22302382.
22. Campochiaro PA, Brown DM, Awh CC, Lee SY, Gray S et al. (2011) Sustained benefits from ranibizumab for macular edema following central retinal vein occlusion: twelve-month outcomes of a phase III study. *Ophthalmology* 118: 2041–2049. doi:10.1016/j.ophtha.2011.02.038. PubMed: 21715011.
23. Brown DM, Campochiaro PA, Bhishitkul RB, Ho AC, Gray S et al. (2011) Sustained benefits from ranibizumab for macular edema following branch retinal vein occlusion: 12-month outcomes of a phase III study.

(DOC)

Acknowledgements

We thank Edith Motschall of the University Medical Centre Freiburg for conducting the literature search in MEDLINE, PreMedline and Embase.

Author Contributions

Conceived and designed the experiments: AP NF CI CS. Performed the experiments: AP NF CI JC BJ CS. Analyzed the data: AP NF CI JC BJ CS. Contributed reagents/materials/analysis tools: AP NF CI JC BJ CS. Wrote the manuscript: AP NF CI JC BJ CS.

- Ophthalmology 118: 1594–1602. doi:10.1016/j.ophtha.2011.02.022. PubMed: 21684606.
24. Centre for Reviews and Dissemination (n.d.) Systematic reviews of adverse effects. Systematic Reviews - CRD's guidance for undertaking reviews in health care. York. pp. 177–198
 25. Moher D, Liberati A, Tetzlaff J, Altman DG (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLOS Med 6: e1000097. doi:10.1371/journal.pmed.1000097.
 26. Epstein DL, Algvere PV, Von Wendt G, Seregard S, Kvant A (2012) Benefit from bevacizumab for macular edema in central retinal vein occlusion: twelve-month results of a prospective, randomized study. Ophthalmology 119: 2587–2591. doi:10.1016/j.ophtha.2012.06.037. PubMed: 22902212.
 27. Parodi MB, Iacono P, Ravalico G (2008) Intravitreal triamcinolone acetate combined with subthreshold grid laser treatment for macular oedema in branch retinal vein occlusion: a pilot study. Br J Ophthalmol 92: 1046–1050. doi:10.1136/bjo.2007.128025. PubMed: 18556425.
 28. Russo V, Barone A, Conte E, Prascina F, Stella A et al. (2009) Bevacizumab compared with macular laser grid photocoagulation for cystoid macular edema in branch retinal vein occlusion. Retina (Phila, Pa) 29: 511–515. doi:10.1097/IAE.0b013e318195ca65. PubMed: 19174717.
 29. Donati S, Barosi P, Bianchi M, Al Oum M, Azzolini C (2012) Combined intravitreal bevacizumab and grid laser photocoagulation for macular edema secondary to branch retinal vein occlusion. Eur J Ophthalmol 22: 607–614. doi:10.5301/ejo.5000085. PubMed: 22101497.
 30. Heier JS, Campochiaro PA, Yau L, Li Z, Saroj N et al. (2012) Ranibizumab for macular edema due to retinal vein occlusions: long-term follow-up in the HORIZON trial. Ophthalmology 119: 802–809. doi:10.1016/j.ophtha.2011.12.005. PubMed: 22301066.
 31. Epstein DLJ, Algvere PV, Von Wendt G, Seregard S, Kvant A (2012) Bevacizumab for macular edema in central retinal vein occlusion: a prospective, randomized, double-masked clinical study. Ophthalmology 119: 1184–1189. doi:10.1016/j.ophtha.2012.01.022. PubMed: 22424833.
 32. Chan CK, Ip MS, VanVeldhuisen PC, Oden NL, Scott IU et al. (2011) SCORE Study report 11: incidences of neovascular events in eyes with retinal vein occlusion. Ophthalmology 118: 1364–1372. doi:10.1016/j.ophtha.2010.11.020. PubMed: 21440942.
 33. Holz FG, Roider J, Ogura Y, Korobelnik J-F, Simader C et al. (2013) VEGF Trap-Eye for macular oedema secondary to central retinal vein occlusion: 6-month results of the phase III GALILEO study. Br J Ophthalmol, 97: 278–84. doi:10.1136/bjophthalmol-2012-301504. PubMed: 23298885.
 34. Ip MS, Oden NL, Scott IU, VanVeldhuisen PC, Blodi BA et al. (2009) SCORE Study report 3: study design and baseline characteristics. Ophthalmology 116: 1770–1777.e1. doi:10.1016/j.ophtha.2009.03.022. PubMed: 19619896.
 35. Schmucker C, Ehlken C, Agostini HT, Antes G, Ruecker G et al. (2012) A safety review and meta-analyses of bevacizumab and ranibizumab: off-label versus goldstandard. PLOS ONE 7: e42701. doi:10.1371/journal.pone.0042701. PubMed: 22880086.
 36. Boyer D, Heier J, Brown DM, Clark WL, Vitti R et al. (2012) Vascular endothelial growth factor Trap-Eye for macular edema secondary to central retinal vein occlusion: six-month results of the phase 3 COPERNICUS study. Ophthalmology 119: 1024–1032. doi:10.1016/j.ophtha.2012.01.042. PubMed: 22440275.
 37. Cugati S, Wang JJ, Rochtchina E, Mitchell P (2006) Ten-year incidence of retinal vein occlusion in an older population: the Blue Mountains Eye Study. Arch Ophthalmol 124: 726–732. doi:10.1001/archophth.124.5.726. PubMed: 16682596.
 38. Klein R, Peto T, Bird A, Vannewkirk MR (2004) The epidemiology of age-related macular degeneration. Am J Ophthalmol 137: 486–495. doi:10.1016/j.ajo.2003.11.069. PubMed: 15013873.
 39. Mity D, Bunce C, Charteris D (2013) Anti-vascular endothelial growth factor for macular oedema secondary to branch retinal vein occlusion. Cochrane Database Syst Rev 1: CD009510: CD009510. Retrieved on published at whilst December year 1111 from doi: 10.1002/14651858.CD009510.pub2. PubMed: 23440840.
 40. Braithwaite T, Nanji AA, Greenberg PB (2010) Anti-vascular endothelial growth factor for macular edema secondary to central retinal vein occlusion. Cochrane Database Syst Rev CD: 007325. Retrieved on published at whilst December year 1111 from doi: 10.1002/14651858.CD007325.pub2 PubMed: 20927757.
 41. Smiddy WE (2011) Economic considerations of macular edema therapies. Ophthalmology 118: 1827–1833. doi:10.1016/j.ophtha.2010.12.034. PubMed: 21507488.
 42. Schmucker C, Ehlken C, Hansen LL, Antes G, Agostini HT et al. (2010) Intravitreal bevacizumab (Avastin) vs. ranibizumab (Lucentis) for the treatment of age-related macular degeneration: a systematic review. Curr Opin Ophthalmol 21: 218–226. doi:10.1097/ICU.0b013e3283386783. PubMed: 20393293.

Intravitreal steroids versus observation for macular edema secondary to central retinal vein occlusion (Review)

Gewaily D, Greenberg PB



**THE COCHRANE
COLLABORATION®**

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2010, Issue 7

<http://www.thecochranelibrary.com>



Intravitreal steroids versus observation for macular edema secondary to central retinal vein occlusion (Review)
Copyright © 2010 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	2
OBJECTIVES	3
METHODS	3
RESULTS	5
DISCUSSION	13
AUTHORS' CONCLUSIONS	19
ACKNOWLEDGEMENTS	19
REFERENCES	19
CHARACTERISTICS OF STUDIES	22
DATA AND ANALYSES	24
APPENDICES	24
WHAT'S NEW	26
HISTORY	26
CONTRIBUTIONS OF AUTHORS	26
DECLARATIONS OF INTEREST	27
SOURCES OF SUPPORT	27
NOTES	28
INDEX TERMS	28

[Intervention Review]

Intravitreal steroids versus observation for macular edema secondary to central retinal vein occlusion

Dina Gewaily¹, Paul B Greenberg¹

¹Division of Ophthalmology, Warren Alpert Medical School of Brown University, Providence, Rhode Island, USA

Contact address: Paul B Greenberg, Division of Ophthalmology, Warren Alpert Medical School of Brown University, Providence, Rhode Island, 02908, USA. paul_greenberg@brown.edu.

Editorial group: Cochrane Eyes and Vision Group.

Publication status and date: Edited (no change to conclusions), published in Issue 7, 2010.

Review content assessed as up-to-date: 9 February 2009.

Citation: Gewaily D, Greenberg PB. Intravitreal steroids versus observation for macular edema secondary to central retinal vein occlusion. *Cochrane Database of Systematic Reviews* 2009, Issue 1. Art. No.: CD007324. DOI: 10.1002/14651858.CD007324.pub2.

Copyright © 2010 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Central retinal vein occlusion (CRVO) is a common retinal vascular abnormality associated with conditions such as hypertension, diabetes, glaucoma, and a wide variety of hematologic disorders. Macular edema (ME) represents an important vision-threatening complication of CRVO. There is no proven treatment; laser photocoagulation is not effective in treating cystoid macular edema secondary to CRVO. Intravitreal steroids, such as triamcinolone acetonide, have been utilized to treat macular edema stemming from a variety of etiologies and may represent a treatment option for CRVO-ME.

Objectives

The objective of this review was to explore the effectiveness and safety of intravitreal steroids in the treatment of CRVO-ME.

Search strategy

We searched the Cochrane Central Register of Controlled Trials (*The Cochrane Library* Issue 4, 2008), MEDLINE (January 1950 to November 2008) and EMBASE (January 1980 to November 2008). There were no language or date restrictions in the search for trials. The electronic databases were last searched on 5 November 2008. For all included primary studies, we used The Science Citation Index and manually reviewed reference lists to identify other possible relevant trials. We contacted researchers in the field, currently working on a randomized controlled trial (RCT) on this topic (The Standard Care versus Corticosteroid for Retinal Vein Occlusion - SCORE study), for information on additional current, past, or unpublished trials.

Selection criteria

We considered RCTs that compared intravitreal steroids of any dosage/duration to observation in the treatment of CRVO-ME for inclusion in this review. We focused on studies that included individuals of any age or gender with unilateral or bilateral disease, with a minimum of six months follow up. Secondly we considered non-randomized studies with the same criteria for description of evidence, however we did not conduct a separate electronic search for finding all non-randomized studies.

Data collection and analysis

We found no RCTs that met the inclusion criteria after independent and duplicate review of the search results.

Main results

We found no relevant RCTs and therefore performed no meta-analysis. Evidence from non-randomized studies is reported in this review.

Authors' conclusions

There is inadequate evidence for the use of intravitreal steroids for CRVO-ME due to a paucity of RCTs and well-designed observational studies on the topic; therefore, it is still an experimental procedure.

PLAIN LANGUAGE SUMMARY

Intravitreal steroids versus observation for macular edema secondary to central retinal vein occlusion

Central retinal vein occlusion (CRVO) is the most common retinal vascular abnormality second to diabetic retinopathy. It classically presents as a unilateral painless loss of vision in individuals over the age of 40 and is associated with conditions such as high blood pressure, diabetes, glaucoma, and hematologic diseases. Macular edema (ME) is a complication of CRVO and is the primary reason for loss of vision in this condition. Currently there is no treatment for this disorder and laser treatment is not effective for CRVO-induced macular edema. Steroid injections in the eye have been used to treat macular edema caused by other eye disorders. This review aimed to examine the risks and benefits of using intravitreal steroids in treating CRVO-ME. While intravitreal steroids can lead to improvements in vision, the effect usually lasts only a few months and there is a risk of developing glaucoma, cataracts, and other complications. Due to an inadequate number of well-designed studies on the topic, we cannot make definitive conclusions and therefore the use of intravitreal steroids is still experimental. More definitive answers may appear when the results of a new randomized controlled trial (the SCORE trial) are published.

BACKGROUND

Description of the condition

Central retinal vein occlusion (CRVO) is a common retinal vascular abnormality associated with conditions such as hypertension, diabetes, glaucoma, and a wide variety of hematologic disorders (AAO 2007). Patients are usually aged 40 years or older and report sudden painless loss of vision in one eye upon waking. Funduscopic evaluation typically reveals intraretinal hemorrhages in all four quadrants ("blood and thunder appearance"), dilation, and tortuosity of the retinal veins. A histopathological study of 29 eyes suggests that central vein occlusion results from the formation of a thrombus at the lamina cribrosa (Green 1981).

Macular edema (ME) represents an important vision-threatening complication of CRVO. The mechanism of macular edema formation is presumed to occur from a hypoxic environment in the retina that leads to changes in retinal capillaries, including an increase in capillary permeability and plasma leakage (Ip 2004). Decreased visual acuity results from disruption of photoreceptor function by an edematous and hemorrhagic macula and, in some cases, ischemic retinal damage (Mandelcorn 2007).

Despite these characteristic clinical and ophthalmic findings, there is no proven treatment. Laser photocoagulation is not effective in treating macular edema secondary to CRVO (CVOS 1995). The lack of effective treatment has prompted interest in other treatment modalities, including medical therapy with anticoagulants, fibrinolytics, acetazolamide, isovolemic hemodilution, anti-vascular endothelial growth factor (anti-VEGF), and angiostatic agents. Surgical options include vitrectomy, chorioretinal anastomosis, direct venous cannulation with injection of fibrinolytics, and radial optic neurotomy (Mohamed 2007). However, none of the aforementioned interventions has been proven effective in treating CRVO-induced cystoid macular edema.

Description of the intervention

Intravitreal steroids, such as triamcinolone acetonide, have been utilized to treat macular edema stemming from a variety of etiologies, including retinal vein occlusion, diabetic retinopathy, uveitis, pseudophakic cystoid macular edema, and exudative macular degeneration (Antcliff 2001; Bashshur 2004; Conway 2003; Jonas 2005). While past attempts using topical or systemic steroids failed

Issue date: July 2011

Dexamethasone intravitreal implant for the treatment of macular oedema secondary to retinal vein occlusion

**This guidance was developed using the
single technology appraisal process**

NICE technology appraisal guidance 229
Dexamethasone intravitreal implant for the treatment of macular oedema secondary to retinal vein occlusion

Ordering information

You can download the following documents from
www.nice.org.uk/guidance/TA229

- The NICE guidance (this document).
- A quick reference guide – the recommendations.
- 'Understanding NICE guidance' – a summary for patients and carers.
- Details of all the evidence that was looked at and other background information.

For printed copies of the quick reference guide or 'Understanding NICE guidance', phone NICE publications on 0845 003 7783 or email publications@nice.org.uk and quote:

- N2616 (quick reference guide)
- N2617 ('Understanding NICE guidance').

This guidance represents the view of NICE, which was arrived at after careful consideration of the evidence available. Healthcare professionals are expected to take it fully into account when exercising their clinical judgement. However, the guidance does not override the individual responsibility of healthcare professionals to make decisions appropriate to the circumstances of the individual patient, in consultation with the patient and/or guardian or carer.

Implementation of this guidance is the responsibility of local commissioners and/or providers. Commissioners and providers are reminded that it is their responsibility to implement the guidance, in their local context, in light of their duties to avoid unlawful discrimination and to have regard to promoting equality of opportunity. Nothing in this guidance should be interpreted in a way which would be inconsistent with compliance with those duties.

National Institute for Health and Clinical Excellence

MidCity Place
71 High Holborn
London WC1V 6NA

www.nice.org.uk

Contents

1	Guidance	4
2	The technology	4
3	The manufacturer's submission	5
4	Consideration of the evidence	19
5	Implementation	37
6	Related NICE guidance	37
7	Proposed date for review of guidance	38
	Appendix A: Appraisal Committee members, and NICE project team	39
	Appendix B: Sources of evidence considered by the Committee	43
	Appendix B: Sources of evidence considered by the Committee	43



NHS Evidence has accredited the process used by the Centre for Health Technology Evaluation at NICE to produce guidance. Accreditation is valid for 3 years from April 2010 and is applicable to guidance produced using the processes described in NICE's 'Guide to the single technology appraisal process' (2009). More information on accreditation can be viewed at www.evidence.nhs.uk

1 Guidance

- 1.1 Dexamethasone intravitreal implant is recommended as an option for the treatment of macular oedema following central retinal vein occlusion.
- 1.2 Dexamethasone intravitreal implant is recommended as an option for the treatment of macular oedema following branch retinal vein occlusion when:
- treatment with laser photocoagulation has not been beneficial,
or
 - treatment with laser photocoagulation is not considered suitable because of the extent of macular haemorrhage.
- 1.3 People currently receiving dexamethasone intravitreal implant for the treatment of macular oedema secondary to branch retinal vein occlusion who do not meet the criteria specified in 1.2 above should have the option to continue treatment until they and their clinicians consider it appropriate to stop.

2 The technology

- 2.1 Dexamethasone intravitreal implant (Ozurdex, Allergan) incorporates a potent corticosteroid that suppresses inflammation in the eye by inhibiting oedema, fibrin deposition, capillary leakage and phagocytic migration. Corticosteroids inhibit the expression of vascular endothelial growth factor (VEGF), a cytokine that is expressed at increased concentrations in macular oedema and is a potent promoter of vascular permeability. Corticosteroids also prevent the release of prostaglandins, some of which are mediators of cystoid macular oedema.
- 2.2 Dexamethasone intravitreal implant has a marketing authorisation for the treatment of adult patients with macular oedema following

either branch retinal vein occlusion (BRVO) or central retinal vein occlusion (CRVO).

- 2.3 The most common adverse reactions are increased intraocular pressure and conjunctival haemorrhage. Conjunctival haemorrhage is related to the intravitreal injection procedure rather than the dexamethasone implant. Other common adverse events include ocular hypertension, vitreous detachment, cataract, subcapsular cataract, vitreous haemorrhage, visual disturbance, vitreous opacities, eye pain, photopsia, conjunctival oedema, and conjunctival hyperaemia. For full details of side effects and contraindications, see the summary of product characteristics.
- 2.4 The cost of a 700-microgram implant and applicator is £870.00 (British National Formulary [BNF] edition 61), excluding VAT. One dexamethasone intravitreal implant is administered usually every 6 months in the affected eye and up to six implants may be given. Costs may vary in different settings because of negotiated procurement discounts.

3 The manufacturer's submission

The Appraisal Committee (appendix A) considered evidence submitted by the manufacturer of dexamethasone and a review of this submission by the Evidence Review Group (ERG; appendix B).

- 3.1 The manufacturer originally submitted evidence of clinical and cost effectiveness for dexamethasone versus best supportive care (observation). Following the consultation after the first Appraisal Committee meeting, the manufacturer submitted a comparison versus bevacizumab. No comparison was made with triamcinolone (analog formulation), which was defined in the scope as a comparator for the treatment of macular oedema following both BRVO and CRVO. Similarly, dexamethasone was not compared with grid laser photocoagulation for non-ischaemic BRVO.

3.2 In the manufacturer's original submission, evidence of clinical effectiveness versus observation was based on two identical randomised, sham-controlled, three-arm parallel-group studies of dexamethasone intravitreal implant in people with macular oedema secondary to BRVO or CRVO. Both studies (GENEVA 008 and GENEVA 009) had an initial 6-month treatment period followed by a 6-month open-label extension in which all patients in both arms of the study who met the re-treatment criteria received a dexamethasone implant. Patients were re-treated if best corrected visual acuity (BCVA) was less than 84 letters or the retinal thickness by optical coherence tomography was more than 250 μ m in the central 1 mm macular subfield and, in the investigators' opinion, the procedure would not put the patient at significant risk. All participants had macular oedema for 6 weeks to 12 months before study entry. Participants were allocated in a 1:1:1 ratio to receive a 700-microgram dexamethasone intravitreal implant (n = 427), a sham implant (n = 426) or a 350-microgram dexamethasone implant (n = 414). This appraisal considered the 700-microgram dose which is the only dose which has a UK marketing authorisation. The sham group had a needleless applicator pressed against the conjunctiva actuated with a click. Investigators were masked to study treatment. The results were presented separately for people with retinal vein occlusion (RVO) and the subgroups of people with macular oedema secondary to CRVO, BRVO, BRVO with macular haemorrhage and BRVO with previous laser treatment. People with BRVO who are eligible for laser therapy were not included as a subgroup. The results from the two studies (GENEVA 008 and GENEVA 009) were pooled and this formed the basis of the evidence considered by the Committee, although the data were also available separately for each study in the manufacturer's submission.

3.3 The results of the pooled analysis showed that for the total RVO population 21.3% of the 427 patients in the intention-to-treat

population receiving dexamethasone had an improvement in BCVA from baseline of at least 15 letters at day 30 compared with 7.5% of 426 patients in the sham group. This rose to 29.3% at day 60 (compared with 11.3% in the sham group) but returned to 21.8% and 21.5% at day 90 and day 180 respectively (compared with 13.1% and 17.6% in the sham group). The differences were statistically significant at day 30 ($p \leq 0.001$), 60 ($p \leq 0.001$) and 90 ($p = 0.008$) but not at day 180 ($p \geq 0.05$). The results for patients who were re-treated at day 180 were presented as academic-in-confidence information and are therefore not presented here.

3.4 For the CRVO subgroup, 21.3% of patients in the dexamethasone group had an improvement in BCVA from baseline of at least 15 letters compared with 6.8% in the sham group at day 30 ($p \leq 0.001$). At day 60, 28.7% in the dexamethasone group had an improvement in BCVA of at least 15 letters compared with 8.8% in the sham group ($p \leq 0.001$). There was no statistically significant difference between the groups at days 90 and 180.

3.5 In the subgroup with BRVO, 21.3% of patients receiving dexamethasone had an improvement in BCVA from baseline of at least 15 letters at day 30 compared with 7.9% in the sham group ($p < 0.001$). The corresponding figures for the subgroup with BRVO with macular haemorrhage were 22.0 % and 8.8 % respectively ($p \leq 0.001$). Both subgroups had statistically significant differences between patients treated with dexamethasone and the sham group at days 60 and 90, but not at day 180. In the subgroup with BRVO and previous laser therapy 22.2% had an improvement in BCVA from baseline of at least 15 letters at day 30 compared with 2.8% in the sham group ($p = 0.028$). Differences between the dexamethasone and sham groups were also statistically significant at days 60 ($p \leq 0.001$), 90 ($p = 0.011$) and 180 ($p = 0.022$).

3.6 The cumulative response rate for time to achieve an improvement in BCVA of at least 15 letters from baseline in the study eye was

statistically significant for dexamethasone versus sham. The difference in mean change from baseline BCVA, the categorical change from baseline BCVA and proportion of patients with an improvement in BCVA of at least 10 letters from baseline in the study eye were statistically significantly higher for dexamethasone versus sham at days 30, 60, 90 and 180 in the pooled analysis. For all RVO at 180 days, the most common adverse events were raised intraocular pressure, eye pain and ocular hypertension. Intraocular pressure was raised in 25.2 % of patients treated with dexamethasone compared with 1.2 % in the sham group. Of patients treated with dexamethasone, 7.4% had eye pain compared with 3.8% in the sham group. Ocular hypertension was experienced by 4% of patients in the treated group compared with 0.7% in the sham group. Presence of anterior chamber cells and retinal neovascularisation were also reported. Other reported adverse events were retinal detachment, retinal tears and cataract. Safety data for the re-treated population (receiving a second implant by day 180) were presented as academic-in-confidence information and are therefore not presented here.

- 3.7 Following the consultation after the first Appraisal Committee meeting, the manufacturer submitted a mixed-treatment comparison of dexamethasone versus bevacizumab. The network of evidence from a systematic review was for BRVO only and included the BRVO data from the GENEVA trials, a non-randomised study by the Branch Vein Occlusion Study group comparing laser with sham treatment (n = 78) and a randomised study by Russo et al. comparing laser with bevacizumab (n = 30). The outcome was improvement in BCVA using standard effect sizes and a fixed-effects model. Dexamethasone was less effective than bevacizumab with a difference in BCVA of –1.74 letters (95% confidence interval [CI] –9.57 to 6.19) when assessed at day 180. However dexamethasone was more effective than bevacizumab

when assessed at day 60, with a gain of 2.55 letters (95% CI –5.28 to 10.48).

- 3.8 For evidence of cost effectiveness, the manufacturer submitted a de novo Markov model that compared treatment with dexamethasone with sham injection in people with macular oedema and vision loss following CRVO or BRVO. Treatment was modelled over a lifetime horizon based on the transition of people between five health states based on the Early Treatment Diabetic Retinopathy Study (EDTRS) measurement of BCVA in the affected eye and death. The worst health state represented visual acuity less than or equal to 38 letters, which equated to severe visual impairment. The best health state was visual acuity of 69 letters or over. The mean BCVA of people in the model was 54 letters, which equates to the second best health state. The patient population was based on data from the GENEVA trials. The model assumed that 90% of people would present with macular oedema in the 'worse-seeing' eye. The model had a cycle length of 1 month for the first 3 months following presentation with RVO, followed by a 3-month cycle in months 4–6 and 6-monthly cycles thereafter. Patients entering the model received dexamethasone or observation. Up to 12 months, transition probabilities were based on pooled patient-level data from the GENEVA studies, including the open-label extension. Beyond 12 months, data were extrapolated from 6- to 12-month data for treatment and re-treatment and 3- to 6-month data for sham. Treatment duration was assumed to last for 2.5 years in people with BRVO and 3 years in people with CRVO; thereafter visual acuity was assumed to be stable.
- 3.9 The data inputs for the manufacturer's model included utility values estimated using the Visual Function Questionnaire Utility Index (VFQ-UI) and mapped onto the health states using an algorithm from a study eliciting preferences from the general population.

Resource use was identified from a systematic review of the literature and input from clinical specialists. Costs included drug cost and medical resource use (hospital visits, monitoring, costs associated with blindness and the cost of treating adverse events, including raised intraocular pressure, cataracts, retinal tears/detachment). Costs associated with treating adverse events were assumed to increase with the third and fourth treatment.

3.10 Key assumptions of the economic model included:

- 90% of people treated would have macular oedema in the 'worse-seeing' eye
- the stabilisation of visual acuity for 2.5 years in people with BRVO and 3 years in people with CRVO
- re-treatment at 6-monthly intervals with a maximum of five injections for BRVO and six injections for CRVO (with assumptions over the number of treatments received)
- a risk of involvement of the other eye of 6.5% in the first year (for those with initial RVO in their 'worse-seeing' eye)
- blindness and an excess mortality hazard of 1.54 associated with a BCVA in the 'better-seeing' eye of 38 or fewer letters (measured by the EDTRS).

Sensitivity analyses included varying utility estimates, costs, stabilisation of visual acuity at day 360, extrapolation assumptions, mortality, involvement of the other eye, discounting, re-treatment, and people with a worse BCVA on entering the model. Results were presented for the entire RVO population and the subgroups of CRVO, BRVO with macular haemorrhage, BRVO with previous laser therapy, BRVO with a diagnosis of 90 days or less at the time of treatment, and BRVO with a diagnosis of more than 90 days at the time of treatment. In the original sensitivity analyses (manufacturer original submission), the factors having the largest impact on estimates of cost effectiveness for the total population were costs associated with vision loss (costs of residential care and

the uptake of residential care), affected eye (proportion of people treated for macular oedema in the 'worse-seeing' eye) and rates of discount.

- 3.11 Following the consultation after the first Appraisal Committee meeting, the model was amended by the manufacturer, resulting in a revised base-case ICER. The revised base case used a ratio of 25% outpatient versus 75% day case procedures for administration and an adjustment to the way that the costs of vision loss were applied to the 'better-seeing' eye in patients whose BCVA in the affected eye falls below $\leq 20/200$. An adjustment of 25% plus a further 10% uplift applied every 6 months was made to the average annual costs associated with severe visual impairment. The model continued to use only the last 3 months of observation data from the GENEVA studies.
- 3.12 In the revised base case for all RVO, the total incremental cost was $\pounds 3698$ for dexamethasone compared with observation and the incremental $\pounds$ ALs were 0.21. The incremental cost-effectiveness ratio (ICER) was $\pounds 17,558$ per $\pounds$ AL gained for dexamethasone intravitreal implant compared with observation in all RVO. In the revised base case for CRVO, the total incremental cost was $\pounds 4732$ and the incremental $\pounds$ ALs were 0.29. The cost per $\pounds$ AL gained was $\pounds 16,522$ for dexamethasone compared with observation. In the revised base case for BRVO with macular haemorrhage, the total incremental cost was $\pounds 3119$ and the incremental $\pounds$ ALs were 0.18. The incremental cost per $\pounds$ AL gained was $\pounds 17,741$. In the revised base case for BRVO with previous laser therapy, the total incremental cost was $\pounds 1857$ and the incremental $\pounds$ ALs were 0.29. The incremental cost per $\pounds$ AL was $\pounds 6361$ for dexamethasone compared with observation for BRVO with previous laser treatment.
- 3.13 In addition, the manufacturer also provided additional scenario analyses in response to requests in the appraisal consultation document. Alternative scenario analyses for the re-treatment rate

included scenarios in which proportions re-treated were as at day 180 for the five injections after the first injection in people with CRVO, proportions re-treated were as at day 180 for the four injections after the first injection in people with BRVO and proportions re-treated were varied between the two extremes of the base case and the GENEVA studies. When proportions re-treated were as at day 180, the ICERs for CRVO, BRVO with macular haemorrhage and BRVO with previous laser treatment increased from £16,522, £17,741 and £6361 to £22,083, £45,878 and £16,548 per £AL gained respectively. When the mid-point was used, the ICERs for CRVO, BRVO with macular haemorrhage and BRVO with previous laser treatment increased from £16,522, £17,741 and £6361 to £20,257, £31,123 and £10,876 per £AL gained respectively.

- 3.14 Following the consultation after the first Appraisal Committee meeting, the manufacturer submitted an additional cost–utility analysis versus bevacizumab for BRVO. This used the mixed-treatment analysis to obtain treatment effects for bevacizumab compared with dexamethasone (a non-significant mean gain of 1.74 letters for bevacizumab compared with dexamethasone at 6 months). Compared with bevacizumab, in the base case for all RVO, the net marginal benefit for dexamethasone (where net marginal benefit is larger than zero and treatment with dexamethasone considered cost effective) was £1927 at £30,000 per £AL gained. The manufacturer submitted a cost-minimisation analysis for dexamethasone versus bevacizumab in CRVO assuming equivalent efficacy for bevacizumab and dexamethasone in CRVO. This was associated with a cost saving of £4463 with dexamethasone which was mainly a result of a lower frequency of injections with dexamethasone compared with bevacizumab and fewer subsequent follow-up visits.

- 3.15 The manufacturer also submitted a second exploratory cost-minimisation analysis for BRVO and a comparison using ranibizumab as a substitute for bevacizumab. This analysis was submitted as commercial in confidence because of its exploratory nature and so is not presented here.
- 3.16 The manufacturer also provided more information about the location and extent of macular haemorrhage in the subgroup of patients for whom laser treatment was not considered appropriate because of macular haemorrhage. In the GENEVA trials, the location and extent was assessed by standardised fundus photographs evaluated by trained graders using a standardised grading protocol (the ETDRS macular oedema grading protocol) and masked to patient group. Retinal haemorrhage in the macula was defined when the grader was at least 90% certain that the retinal haemorrhage was present in the grid (macular area).
- 3.17 The ERG considered the GENEVA trials to be of high quality. Although there was a statistically significant increase in the BCVA based on the mean letter score with the dexamethasone implant, the ERG did not consider this to be clinically significant because most patients did not achieve a 15-letter improvement from baseline. However, a higher proportion had an improvement of at least 10 letters. The effectiveness of the dexamethasone implant appeared to peak at around 60 days. The ERG highlighted that the trial protocol did not allow for early re-treatment and during the trial and open-label follow-on patients received only two injections of dexamethasone. The ERG noted that the main benefit from re-treatment was in patients whose condition had responded during the initial 180-day trial period. The ERG also commented that the number of treatments needed in practice is not known and that clinical opinion estimated a maximum of six.
- 3.18 The ERG also highlighted that because the trial included a maximum of two dexamethasone treatments, the impact of up to

six treatments on the incidence of adverse events was not known. The ERG also expressed concern over the size of the needle which is larger for dexamethasone intravitreal implant than for other treatments. The ERG stated that the main weaknesses in the evidence were lack of long-term follow-up data and data on earlier re-treatment before 180 days.

- 3.19 The ERG considered the robustness of the manufacturer's original model. It highlighted cost inputs (particularly the cost of dexamethasone administration and the cost of severe visual impairment), structural assumptions in the model (such as the duration of trial data on which to base extrapolation of health states in the treatment and observation arms, assumptions related to the stability of visual acuity in 'resolved' patients, the proportion of people who will present with RVO in their 'worse-seeing' eye and the modelling of fellow eye involvement).
- 3.20 The ERG considered that a number of the unit costs applied in the model had been overestimated. For example, the cost of administering the dexamethasone intravitreal implant might have been overestimated because the implant could be given on an outpatient basis (£150) but costs were based on day-case (£648) care in the manufacturer's submission. The ERG also estimated that the cost of residential care was £16,999 instead of £23,972 as used in the base case, and the cost of cataract extraction was £789 rather than £965 as in the base-case model.
- 3.21 The ERG conducted a sensitivity analysis on the unit costs applied in the manufacturer's model and other assumptions related to the extrapolation of effectiveness data beyond the trial. According to the ERG, key uncertainties related to the extrapolation of data remain in the evaluation of cost effectiveness. The likely maximum number of dexamethasone administrations and frequency of re-treatment, the likelihood of resolution, the likelihood of cataract development and extraction, the likelihood of involvement of the

other eye and the likelihood of the RVO leading to macular oedema are all important aspects of this uncertainty.

3.22 The ERG highlighted factors in the manufacturer's original sensitivity analyses that appeared to have a particular impact on cost effectiveness. When visual acuity was assumed to be stable after a year with no further dexamethasone treatments (rather than at 2.5 years with BRVO and 3 years with CRVO) the base-case ICER increased from £7368 to £10,764 per £AL gained for the RVO population. When those not treated were all assumed to have the observation transition probabilities applied up to 2.5 years for BRVO and 3 years for CRVO (rather than transition probabilities weighted by proportion of people who were not treated who resolved at day 180 and those who discontinued treatment for other reasons), the base-case ICER increased from £7368 to £24,924 per £AL gained for the RVO population. When the proportions re-treated were based on the re-treatment rate in the trial (day 180) for the five injections after the first in CRVO (85.7%) and the four injections after the first in BRVO (78.8%), this increased the base-case ICER from £7368 to £19,100 per £AL gained for the RVO population. When there was a decline in vision in 1.5% of patients in each health state, worsening by one health state every 6 months was assumed (compared with visual stability from 2.5 years for BRVO and 3 years for CRVO in the base case). This had a small effect on the base-case ICER, which increased from £7368 to £7685 per £AL gained for the RVO population.

3.23 The ERG also conducted its own additional exploratory analyses on the manufacturer's original economic model which updated the modelling of fellow eye involvement. The ERG compared the manufacturer's model, which included the Weibull function for fellow eye involvement, with the same model with no fellow eye involvement. At the same time, the ERG also varied the proportion of people entering the model with macular oedema in the

'worse-seeing' eye from 90% (as in the manufacturer's model) to 97% (based on the proportion treated in the 'worse-seeing' eye in the trials). The alternative assumption of no fellow eye involvement changed the base-case ICERs for CRVO, BRVO with macular haemorrhage and BRVO with previous laser from £6041, £7987 and dominant to £17,279, £34,277 and £11,905 per £AL£ gained respectively. The alternative assumption of 97% with macular oedema in the 'worse-seeing' eye changed the base-case ICER for CRVO, BRVO with macular haemorrhage and BRVO with previous laser from £6041, £7987 and dominant to £15,800, £10,206 and dominant per £AL£ gained respectively.

- 3.24 In addition, the ERG questioned the way in which 6-month data from the open-label phase were used for the extrapolation of results with dexamethasone treatment and the use of 3- to 6-month data from the trial phase for extrapolation in the observation arm of the original model. After the first Committee meeting the Committee had requested that all data be included in a revised model. When extrapolation was based on 6- to 12-month data from the open-label phase of the trial for dexamethasone treatment and 0- to 6-month data for the observation arm (with 90% of people being treated in the 'worse-seeing' eye) the base-case ICER increased from £6041 to £15,395 per £AL£ gained for the RVO population.
- 3.25 The manufacturer submitted two additional original base-case models: one involving structural changes in the modelling of fellow eye involvement, and another which also included lower costs of dexamethasone administration. The revised modelling of fellow eye involvement took account of age with differential survival by collapsing the model into two health states. Without inclusion of lower costs for dexamethasone administration, the revised model was associated with higher ICERs for dexamethasone versus observation of £10,271 per £AL£ gained for all RVO, £8165 for

CRVO, £11,403 for BRVO with macular haemorrhage, and dominance for dexamethasone over observation for BRVO with previous laser treatment. With revised cost assumptions (outpatient appointments rather than day-case costs, reduced costs of residential care and reduced costs of cataract removal), ICERs were reduced to £7616 per £AL gained for all RVO, £6221 for CRVO, £8848 for all BRVO, £8313 for BRVO with macular haemorrhage, and dominance for dexamethasone over observation for BRVO with previous laser treatment.

3.26 The ERG noted that the manufacturer's model applied the cost associated with severe visual impairment to people with visual acuity less than or equal to 38 letters in the 'worse-seeing' eye. The ERG noted that the cost of severe visual impairment should only be applied when both eyes have visual acuity less than 38 letters. The ERG conducted a sensitivity analysis in which the cost of severe visual impairment was applied when both eyes entered the worse state (visual acuity less than 38 letters) and assumed perfect correlation in the BCVA of both eyes and no correlation between the eyes. This increased the base-case ICER for CRVO from £6221 to £15,956 per £AL gained when there was perfect correlation and to £18,091 per £AL gained when there was no correlation between eyes. The corresponding base-case ICERs for BRVO with macular haemorrhage increased from £8313 to £9674 and £21,443 per £AL gained respectively.

3.27 Following the consultation after the first Appraisal Committee meeting, the ERG noted that the manufacturer's revised analysis of cost effectiveness included several deviations from the analysis requested by the Committee at its first meeting. The manufacturer included administration costs for 75:25 day-case to outpatient procedures whereas the Committee had requested day-case costs for all patients. No bevacizumab vial sharing was included in the base case. Extrapolation from data for the last 3 months of

observation in the GENEVA studies was used although the Committee requested that all 6 months of data be used. In addition, severe visual impairment was not modelled as requiring bilateral involvement. Cost of reduction of severe visual impairment was reduced by 25% as a proxy with 10% applied every 6 months. The ERG noted that the 10% was not applied annually and it was unclear from where the 10% figure was derived. When the ERG removed this 10% uplift per cycle, the manufacturer's estimated ICERs for CRVO, BRVO with macular haemorrhage and BRVO with previous laser treatment increased from £16,522, £17,741 and £6361 to £22,831, £23,847 and £12,857 per £AL gained respectively. When the ERG included all 6 months of observational data from the GENEVA studies, the manufacturer's estimated ICERs for CRVO, BRVO with macular haemorrhage and BRVO with previous laser treatment increased from £16,522, £17,741 and £6361 to £25,336, £38,489 and £11,650 per £AL gained respectively.

- 3.28 The ERG considered that a number of the unit costs applied in the revised model had been overestimated for bevacizumab and the costs used favoured dexamethasone. For example, administration costs were used for 25% outpatient and 75% day-case procedures, there was uncertainty over the calculation of the administration costs for outpatients, a high bevacizumab cost was used and a large number of bevacizumab administrations included. Also, the ERG stated that the frequency of cataracts with steroids was underestimated. When the ERG made adjustments for these costs (a cost of bevacizumab of £50, less frequent bevacizumab administration and follow-up based on the Pan American Collaborative Retina Study group data, and 100% outpatient procedures), the base-case cost minimisation for CRVO, which estimated a cost saving of £4463 with dexamethasone versus bevacizumab, became a cost saving for bevacizumab of £2127 versus dexamethasone.

- 3.29 The ERG highlighted that there was a considerable amount of data on bevacizumab treatment from trials including bevacizumab, laser therapy and triamcinolone in RVO (992 observed patient-years) and agreed that this type of observational and uncontrolled data generates uncertainty. The ERG concluded that more data could have been included in the network model by using data on triamcinolone treatment.
- 3.30 Full details of all the evidence are in the manufacturer's submission and the ERG report, which are available from www.nice.org.uk/guidance/TA229

4 Consideration of the evidence

- 4.1 The Appraisal Committee reviewed the data available on the clinical and cost effectiveness of dexamethasone, having considered evidence on the nature of macular oedema secondary to RVO and the value placed on the benefits of dexamethasone by people with the condition, those who represent them, and clinical specialists. It also took into account the effective use of NHS resources.
- 4.2 The Committee considered the impact of macular oedema on the everyday life of patients. It heard from patient experts about the problems associated with macular oedema and related vision loss, including difficulties with driving in the dark, taking part in hobbies such as craftwork, picking up small things, reading, using computers and distinguishing objects in crowded places. The patient experts also noted the negative impact of macular oedema on social activities. The patient experts acknowledged that although people may be worried about having an injection in the eye, this is preferable to loss of vision. They also noted that the injections were administered from the side and so could not be seen. The Committee concluded that loss of vision caused by macular

oedema secondary to RVO had a negative impact on health-related quality of life and that there was a need for appropriate treatment.

- 4.3 The Committee considered the decision problem submitted by the manufacturer. It noted that the manufacturer's initial submission compared dexamethasone with best supportive care (observation) alone. This was not consistent with the scope, which defined the comparators for both BRVO and CRVO as triamcinolone acetonide (IVTA; $\square$ enalog formulation or equivalent), bevacizumab and best supportive care, and for non-ischaemic BRVO as grid pattern photocoagulation.
- 4.4 The Committee considered the use of triamcinolone to treat RVO in U $\square$ clinical practice. It heard from clinical specialists that the triamcinolone formulation available in the U $\square$ is contraindicated for ocular use ($\square$ enalog formulation); it also heard from the manufacturer of the intraocular formulation of triamcinolone (Trivaris) that it is not available in the U $\square$ and would not be marketed anywhere in the world. The Committee concluded that triamcinolone is not available and would not be a relevant comparator for dexamethasone.
- 4.5 The Committee considered the use of bevacizumab to treat RVO in U $\square$ clinical practice. It noted that bevacizumab does not have a U $\square$ marketing authorisation for the treatment of RVO and heard from patient experts that they were concerned about the use of any unlicensed treatments for which there was no formal post-marketing surveillance, particularly if there were alternatives that had a U $\square$ marketing authorisation. The Committee understood that licensing is not a prerequisite for consideration of a comparator in a NICE appraisal as long as it is in routine use or best practice. It heard from clinical specialists that bevacizumab is currently reasonably widely used in the NHS, being routinely used in some centres, for some RVO cases only in others, and not at all in other centres. It heard from the experts that some clinicians and patients

want more information about its long-term efficacy and safety. It also noted consultee comments following consultation confirming the use of bevacizumab by many ophthalmologists in the UK. The Committee considered the results of the independent survey supplied by the manufacturer which indicated that a proportion of consultant ophthalmologists use bevacizumab for RVO either regularly or occasionally. It acknowledged that the sample size for the survey was small and could be subject to some selection bias. The Committee concluded that bevacizumab is in routine use to treat RVO in some parts of the UK; and it is therefore relevant for consideration as a comparator to dexamethasone.

- 4.6 The Committee considered laser photocoagulation as a treatment for BRVO. It noted that the manufacturer restricted the analysed population to those people with BRVO who cannot receive or have already tried and not benefited from laser photocoagulation. The Committee noted that the manufacturer defined those who cannot receive laser photocoagulation as people with macular haemorrhage. It noted comments from clinical specialists and consultees that all cases of RVO have a degree of macular haemorrhage and the decision to treat using laser photocoagulation is relatively subjective. Therefore, the Committee recognised that for a subgroup of people with BRVO with a lesser extent of macular haemorrhage, laser photocoagulation would be a treatment option and that this subgroup is within the licensed indication for dexamethasone. The Committee concluded that because no evidence had been provided for this subgroup, any recommendation for dexamethasone would be optimised to those subgroups of people with BRVO who cannot receive or have already tried and not benefited from laser photocoagulation.

Clinical effectiveness

- 4.7 The Committee considered the potential of dexamethasone to offer additional health-related benefits compared with currently available

treatment options. It heard from the patient experts about the impact of dexamethasone on their quality of life. Patients advised that after the administration of dexamethasone their sight improved and they were able to resume normal daily activities.

4.8 The Committee considered the evidence for the efficacy of dexamethasone intravitreal implant compared with best supportive care from the GENEVA trials. It noted that the primary outcome was the percentage of patients with an improvement of BCVA of 15 letters or more, which represented a gain of three lines on the EDTRS chart (this enables patients to see letters half the height of those they could see before). The Committee heard from clinical specialists that this represented the gold standard for assessing the effect of treatment on visual acuity, although a gain of 10 letters is also considered to be clinically significant.

4.9 The Committee considered the results from the trials. It first considered the pooled primary outcome data from the GENEVA trials for the entire population with macular oedema following RVO. The Committee noted that dexamethasone is associated with a statistically significant improvement in visual acuity (based on percentage of patients with an improvement of BCVA of 15 letters or more) compared with sham treatment at day 30, 60 and 90, but there was no statistically significant improvement at day 180. The Committee also noted that dexamethasone was associated with a statistically significant improvement in mean change in BCVA at day 30, 60, 90 and 180 for all people with macular oedema following RVO. The Committee also considered the pooled primary outcome data from the GENEVA trials for the CRVO and BRVO subgroups. It noted that dexamethasone was associated with a statistically significantly higher proportion of people gaining BCVA of 15 letters or more compared with sham treatment for CRVO (day 30 and 60), BRVO (day 30, 60 and 90), BRVO with macular haemorrhage (day 30, 60 and 90) and at all time points for BRVO

with previous laser treatment. These results for the subgroups were similar to those obtained for the combined population with RVO. The Committee concluded that treatment with dexamethasone is clinically effective when compared with best supportive care.

4.10 The Committee considered the evidence for adverse events associated with dexamethasone. These included cataracts, raised intraocular pressure and infection. The Committee noted that evidence was limited by trial duration to adverse events after two treatments with data collected up to 360 days. The Committee discussed the impact of dexamethasone in causing cataracts and the potential issues for people with diabetes, who have a higher risk of developing glaucoma and cataracts. The Committee heard from the clinical specialist that a similar infection risk to that of its comparators would be expected. It also heard that the incidence of raised intraocular pressure was 25% with dexamethasone but that this is usually well managed with eye drops and that 1% of people required laser or surgical procedures for management of elevated intraocular pressure. A higher percentage was likely to need surgery for cataracts according to the number of dexamethasone injections given. The Committee concluded that there were some concerns about the long term safety of dexamethasone treatment because the marketing authorisation is based on a evidence base trial with two re-treatments over 360 days and the manufacturer assumed that up to six treatments would be given and there are limited data on long-term treatment and multiple re-treatment (see section 3.6).

4.11 The Committee considered the evidence for re-treatment from the open-label extension of the GENEVA trials (which was submitted as academic-in-confidence information, as described in section 3.6, and is therefore not presented here). The Committee noted that for the whole population the proportion of patients with an improvement in BCVA of at least 15 letters was higher in the group

that received dexamethasone at day 0 and day 180 compared with those who received sham at day 0 and dexamethasone at day 180. The Committee also considered the expected frequency of treatment with dexamethasone at day 180. The Committee heard from the clinical specialist that the criteria for re-treatment were based on the patient's experience with previous treatment (whether their vision had initially improved with treatment but had started to deteriorate), deterioration in visual acuity as assessed by BCVA and the persistence of macular oedema. The Committee was further advised that it is difficult to use a cut-off value for visual acuity that would indicate re-treatment, but that with a loss of five letters it would be appropriate to consider re-treatment. The Committee agreed with the clinical specialist that although the safety data relate to 6-monthly treatment, it is expected that clinicians may re-treat at 4 months in clinical practice, but may not treat more frequently because of the risk of adverse events from the accumulation of dexamethasone in the eye.

- 4.12 The Committee then considered the clinical effectiveness of dexamethasone compared with bevacizumab. It heard from the clinical specialist that bevacizumab is generally considered by ophthalmologists to be efficacious and safe, although there is some uncertainty as to the optimal dosing schedule, relative effectiveness versus dexamethasone and frequency of treatments. The Committee considered the results of the mixed-treatment comparison provided by the manufacturer in response to requests made at the first Appraisal Committee meeting. It noted that the difference in BCVA was marginally favourable for dexamethasone over bevacizumab at day 60, but that at day 180 the difference in BCVA was marginally favourable for bevacizumab over dexamethasone. However, at both time points the confidence intervals were wide and therefore little certainty could be placed in either of the findings. The Committee noted the ERG's misgivings about the lack of completeness of the mixed-treatment comparison

and also that the two cost–utility analyses provided by the manufacturer favoured bevacizumab in terms of efficacy. Nevertheless the Committee considered that a clinical efficacy advantage for dexamethasone over bevacizumab could not be concluded with certainty based on current evidence.

- 4.13 The Committee considered the mode of delivery and adverse events associated with dexamethasone and bevacizumab, including infection and needle phobia. The clinical specialist informed the Committee that dexamethasone was administered with a larger needle than existing treatments and that dexamethasone was associated with increased intraocular pressure and cataracts. However, more injections of bevacizumab were needed and this presented a greater infection risk. The clinical specialist also informed the Committee that with bevacizumab, which had no agreed protocol for use, sterile endophthalmitis had been reported, but was rare. The Committee concluded that the adverse events associated with dexamethasone were expected to be more severe and more difficult to treat than those associated with bevacizumab.

Cost effectiveness

- 4.14 The Committee considered the cost effectiveness of dexamethasone relative to best supportive care. It noted the manufacturer's revised base-case estimate of the ICER for all people with RVO of £17,600 per £AL gained. It understood that this estimate incorporated an assumption that 90% of people would be treated for macular oedema in their 'worst-seeing eye'. It heard from clinical specialists that in practice this proportion would be between 90% and 97%. It further heard from clinical specialists that the manufacturer's estimate of the percentage (6.5%) of people who require treatment for the other eye (fellow eye involvement) was plausible. Patient experts had highlighted the importance of treating the first eye affected, even though overall acuity depends

mostly on the 'better-seeing' eye, because RVO and other eye conditions may later affect the critical second eye. The Committee concluded that in this instance it was appropriate to treat the first eye affected.

- 4.15 The Committee considered the manufacturer's response to requests made at the first Appraisal Committee meeting. It understood that the revised base-case estimate of the ICER incorporated revised assumptions for the setting for the administration of dexamethasone and the cost of care for people with blindness. It further understood that the manufacturer had provided scenario analyses to demonstrate the effect of different re-treatment rates. The Committee noted that the revised base-case ICER did not incorporate the Committee's request that the extrapolation of data for the observation arm should be based on those trial data between 0 and 6 months rather than 3–6 months. The Committee discussed each of these assumptions, and the results of the sensitivity analyses for re-treatment rates, in turn.
- 4.16 The Committee considered the setting for the administration of dexamethasone. It noted that the revised base case was based on an assumption of 75% of procedures being performed in a day-case unit and 25% in an outpatient clinic, instead of 100% in an outpatient clinic as requested by the Committee at the first Appraisal Committee meeting. The Committee noted comments from consultees that it was reasonable to expect a higher proportion to be performed in a day-case unit while ophthalmologists are gaining familiarity with dexamethasone administration, but that in time it would be expected that most would be performed in an outpatient clinic. The Committee therefore accepted the proportions used by the manufacturer in the revised base case for the setting for the administration of dexamethasone.

- 4.17 The Committee considered the revised assumption for the cost associated with blindness in the revised base case. It understood that the manufacturer had accepted the rationale for applying this cost to only those people who have both eyes in the worst health state. However, the Committee noted that in the revised base-case estimate, the cost associated with blindness was reduced as a proxy for adjusting the proportion of people who fall into this health state of severe visual impairment in both eyes. The manufacturer adjusted the cost associated with blindness to 25% of the full cost (that is, £1490 instead of £5963), and increased this every 6 months by 10% of the reduced cost (that is, an increase of £149 for each cycle). The Committee accepted this to be a reasonable reduction and concluded that although it would have been preferable to have adjusted the assumptions around the number of people with severe visual impairment in both eyes, the manufacturer's adjustment went some way towards meeting the Committee's request and was considered acceptable.
- 4.18 The Committee considered the extrapolation of health-state data beyond the trial for the observation arm. It noted the manufacturer's assertion that it was inappropriate to include 0- to 3-month data for BRVO because spontaneous resolution occurs in approximately a quarter of people during that period. The Committee also heard from the manufacturer that although it may be appropriate to include 0- to 3-month data for CRVO, the ERG's exploratory analysis used aggregate data rather than transition probabilities based on individual patient data. The manufacturer stated that this had led to an overestimation of the ICER for CRVO by the ERG. The Committee accepted that spontaneous resolution occurs in BRVO and therefore it was not appropriate to include the first 3 months of data for this subgroup. It further concluded that using individual patient data for the first 6 months may have been relevant for CRVO, but that the increased ICER for CRVO from the ERG's analysis (from £16,500 to £25,300 per £AL gained

including a decrease from 0.29 to 0.22 of net total QALYs and an increase in total cost from £4732 to £5469) may be an overestimate. The Committee concluded that 0- to 3-month data for the observation arm of the GENEVA trials for people with CRVO would not have a substantial impact on the overall ICER and therefore accepted the manufacturer's use of 0- to 3-month data in the revised base case.

4.19 The Committee considered the sensitivity analyses around re-treatment rates provided by the manufacturer in response to requests made by the Committee at the first meeting. The Committee understood that the manufacturer had combined the revised assumptions relating to the setting for administering dexamethasone (see section 4.16) and the costs associated with blindness (see section 4.17) with three scenario analyses for different re-treatment rates. The Committee noted that the re-treatment scenarios covered the extreme positions of both the base-case estimates (in which re-treatment rates were based on clinical opinion) and those re-treated at day 180 in the GENEVA trials. The Committee noted that for all people with RVO, the ICER varied between £17,600 (including incremental costs of £3698 and incremental QALYs of 0.40) and £34,700 (including incremental costs of £8041 and incremental QALYs of 0.44) per QALY gained for these scenarios respectively. The Committee further noted the manufacturer's scenario analysis in which the re-treatment rates were at the mid-points between the two extremes. With this assumption the ICER for all people with RVO was £26,300 per QALY gained. The Committee accepted the assumption of mid-point re-treatment rates.

4.20 On the basis of previous discussions (see sections 4.17 to 4.21), the Committee accepted the revised manufacturer's base-case estimate using the mid-point re-treatment rates from the manufacturer's scenario analysis as the most appropriate estimate

of the ICER for its consideration. It therefore concluded that the decision regarding the cost effectiveness of dexamethasone compared with best supportive care should be based on the manufacturer's ICER of £26,300 per QALY gained for all people with RVO. The Committee further concluded that this represented an acceptable level of cost effectiveness in this case and that dexamethasone intravitreal implant for the treatment of RVO represents a cost-effective use of NHS resources when compared with best supportive care.

4.21 The Committee considered the cost effectiveness of dexamethasone relative to bevacizumab. It understood that the manufacturer had presented a cost-minimisation approach for CRVO and that the cost assumptions from this approach had been used in both cost–utility analyses (that is, the mixed-treatment comparison, and the indirect comparison using a different anti-VEGF treatment as a proxy for bevacizumab) (see sections 3.14 and 3.15). The Committee therefore considered that its discussion on the relative cost effectiveness of dexamethasone compared with bevacizumab should focus on the assumptions within the manufacturer's cost-minimisation analysis.

4.22 The Committee considered the assumptions in the manufacturer's cost-minimisation analysis between dexamethasone and bevacizumab. It noted comments from the ERG that the cost of bevacizumab had been modelled on the higher of the two estimates from the two suppliers of bevacizumab for eye treatment (that is, £150 per vial rather than £50 per vial). The ERG further highlighted that the manufacturer's estimate of the average number of doses of bevacizumab was 13.75 in the first 2 years, but that the Pan American Collaborative Retina Study Group had found an average of 7 doses over this time period. The Committee heard from the manufacturer that this study was a small cohort follow-up study and may not be representative of UK clinical practice. In

addition the manufacturer highlighted to the Committee that the recently reported HORIZON study found that 3.8 injections per year of ranibizumab were insufficient to maintain the clinical effects seen in the first year (ranibizumab, another anti-VEGF, was given monthly in the CRUISE study). Therefore the manufacturer considered 7 doses of bevacizumab over 2 years to be lower than UK clinical practice. The Committee understood that the cost-minimisation analysis also assumed that the setting for the administration of both bevacizumab and dexamethasone would be the same (that is, 75% of procedures would be carried out in a day-case unit and 25% would be carried out in an outpatient clinic). The Committee noted the ERG's view that all anti-VEGF injections are currently administered in an outpatient clinic. The Committee concluded, on the one hand, that the cost of bevacizumab may have been overestimated, and on the other, that there was uncertainty around the number of injections of bevacizumab. The Committee concluded that the assumptions used in the manufacturer's cost-minimisation analysis between dexamethasone and bevacizumab may overestimate the cost of bevacizumab, but considered the cost may not be as low as suggested by the ERG.

- 4.23 The Committee considered the results of the cost-minimisation analysis. It noted that the manufacturer estimated the cost of dexamethasone treatment to be approximately £4500 less per patient than bevacizumab, but that with the ERG's amendments, dexamethasone could cost up to £2100 more per patient than bevacizumab. On the basis of previous discussion (see section 4.22) the Committee accepted that there was some uncertainty around the assumptions used to calculate the cost of bevacizumab and concluded that it was as likely that dexamethasone might cost more than bevacizumab as it was that dexamethasone might cost less. The Committee therefore concluded that a cost advantage of

dexamethasone compared with bevacizumab had not been conclusively demonstrated.

- 4.24 In summary, the Committee considered the costs and clinical effectiveness of dexamethasone. It considered that comparisons with best supportive care and bevacizumab were both relevant to the NHS. It acknowledged that an incremental analysis in which dexamethasone, bevacizumab and best supportive care were simultaneously assessed was not available. The Committee recognised the difficulties within the evidence base for bevacizumab and commended the manufacturer's attempts to provide a comparison of the relative clinical and cost effectiveness of bevacizumab and dexamethasone. It considered that a clinical efficacy advantage for dexamethasone over bevacizumab could not be concluded with certainty based on presently available evidence and that the adverse events associated with dexamethasone were expected to be more severe and more difficult to treat than those associated with bevacizumab. It also accepted that there was uncertainty around the assumptions used in the cost minimisation comparing dexamethasone with bevacizumab. Based on the manufacturer's revised ICER of £26,300 per QALY gained compared with best supportive care for all people with RVO, the Committee concluded that dexamethasone intravitreal implant represents a cost-effective use of NHS resources and should be offered as an option for the treatment of macular oedema secondary to RVO.
- 4.25 The Committee considered whether there were any equalities considerations affecting population groups protected by equality legislation and concluded that there were no equality issues relating to this appraisal that required addressing in the guidance.

Summary of Appraisal Committee's key conclusions

TA229	Appraisal title: Dexamethasone intravitreal implant for the treatment of macular oedema secondary to retinal vein occlusion	Section
Key conclusion		
Dexamethasone intravitreal implant is recommended as an option for the treatment of macular oedema following central retinal vein occlusion.		1.1
Dexamethasone intravitreal implant is recommended as an option for the treatment of macular oedema following branch retinal vein occlusion when:		1.2
• treatment with laser photocoagulation has not been beneficial, or • treatment with laser photocoagulation is not considered suitable because of the extent of macular haemorrhage.		1.3
People currently receiving dexamethasone intravitreal implant for the treatment of macular oedema secondary to retinal vein occlusion should have the option to continue treatment until they and their clinicians consider it appropriate to stop.		
Current practice		
Clinical need of patients, including the availability of alternative treatments	The Committee concluded that triamcinolone (□enalog formulation) is not available and is not a relevant comparator for dexamethasone.	4.4
	The Committee concluded that bevacizumab is in routine use in some parts of the U□ to treat macular oedema secondary to retinal vein occlusion (RVO) and is therefore relevant for consideration as a comparator for dexamethasone.	4.5
	The Committee recognised that for a subgroup of people with branch retinal vein occlusion (BRVO) with a lesser extent of macular haemorrhage, laser photocoagulation would be a treatment option and that this subgroup is within the licensed indication for dexamethasone. The Committee concluded that because no evidence had been provided for this subgroup, any recommendation for dexamethasone would be restricted to those subgroups of people with BRVO who cannot receive or have already tried and not benefited from laser photocoagulation.	4.6
	The Committee concluded that loss of vision caused by macular oedema secondary to RVO had a negative impact on health-related quality of life and that there was a need for appropriate treatment.	4.2
The technology		
Proposed benefits of the technology How innovative is the	Dexamethasone intravitreal implant incorporates a potent corticosteroid that suppresses inflammation in the eye by inhibiting oedema, fibrin deposition,	2.1

technology in its potential to make a significant and substantial impact on health-related benefits□	capillary leakage and phagocytic migration.	
What is the position of the treatment in the pathway of care for the condition□	Dexamethasone intravitreal implant has a marketing authorisation for the treatment of adult patients with macular oedema following either BRVO or central retinal vein occlusion (CRVO). Comparator treatments include best supportive care (observation), bevacizumab, triamcinolone (□enalog formulation) and laser therapy (BRVO only).	2.2 4.3
Adverse effects	The Committee concluded that there were some concerns about long-term safety of dexamethasone treatment because the marketing authorisation is based on an evidence base trial with two re-treatments over 360 days and the manufacturer assumed that up to six treatments would be given and there are limited data on long-term treatment and multiple re-treatment. The Committee concluded that the adverse events associated with dexamethasone were expected to be more severe and more difficult to treat than those associated with bevacizumab.	4.10 4.13
Evidence for clinical effectiveness		
Availability, nature and quality of evidence	The Evidence Review Group (ERG) considered the GENEVA trials, RCTs of dexamethasone versus sham treatment, to be of high quality. The ERG highlighted that there was a considerable amount of data on bevacizumab treatment from trials including bevacizumab, laser therapy and triamcinolone in RVO (992 observed patient-years) and agreed that this type of observational and uncontrolled data generates uncertainty. The Committee recognised the difficulties within the evidence base for bevacizumab and commended the manufacturer's attempts to provide a comparison of the relative clinical and cost effectiveness of bevacizumab and dexamethasone.	3.17 3.29
Relevance to general clinical practice in the NHS	The Committee noted that the primary outcome in the GENEVA trials was the percentage of patients with an improvement of BCVA of 15 letters or more. The Committee heard from clinical specialists that this represented the gold standard for assessing the effect of treatment on visual acuity, although a gain of 10 letters is also considered to be clinically significant.	4.8

Uncertainties generated by the evidence	The Committee considered that a clinical efficacy advantage for dexamethasone over bevacizumab could not be concluded with certainty based on current evidence.	4.12
Are there any clinically relevant subgroups for which there is evidence of differential effectiveness□	The Committee noted that dexamethasone was associated with a statistically significantly higher proportion of people gaining BCVA of 15 letters or more compared with sham treatment for CRVO (day 30 and 60), BRVO (day 30, 60 and 90), BRVO with macular haemorrhage (day 30, 60 and 90) and at all time points for BRVO with previous laser treatment. These results for the subgroups were similar to those obtained for the combined population with RVO.	4.9
Estimate of the size of the clinical effectiveness including strength of supporting evidence	Of the 427 patients in the intention-to-treat population receiving dexamethasone, 21.3% had an improvement in BCVA from baseline of at least 15 letters at day 30 compared with 7.5% of 426 patients in the sham group. The Committee also noted that dexamethasone was associated with a statistically significant improvement in mean change in BCVA at day 30, 60, 90 and 180 for all people with macular oedema following RVO compared with sham. The Committee noted that the difference in BCVA was marginally favourable for dexamethasone over bevacizumab at day 60, but that at day 180 the difference in BCVA was marginally favourable for bevacizumab over dexamethasone. However, at both time points the confidence intervals were wide and therefore little certainty could be placed in either of the findings. The Committee considered that a clinical efficacy advantage for bevacizumab over dexamethasone could not be concluded with certainty based on current evidence.	3.3 4.9 4.12
Evidence for cost effectiveness		
Availability and nature of evidence	The Committee acknowledged that an incremental analysis in which dexamethasone, bevacizumab and best supportive care were simultaneously assessed was not available. The Committee recognised the difficulties within the evidence base for bevacizumab and commended the manufacturer's attempts to provide a comparison of the relative clinical and cost effectiveness of bevacizumab and dexamethasone.	4.24

Uncertainties around and plausibility of assumptions and inputs in the economic model	The Committee considered the setting for the administration of dexamethasone. It noted that the revised base case was based on an assumption of 75% of procedures being performed in a day-case unit and 25% in an outpatient clinic. The Committee accepted the proportions used by the manufacturer in the revised base case.	4.16
	The Committee considered the revised assumption for the cost associated with blindness in the revised base case. The manufacturer adjusted the cost associated with blindness to 25% of the full cost (that is, £1490 instead of £5963), and increased this every 6 months by 10% of the reduced cost (that is, an increase of £149 for each cycle). The Committee accepted this to be a reasonable reduction and concluded that although it would have been preferable to have adjusted the assumptions around the number of people with severe visual impairment in both eyes, the manufacturer's adjustment went some way towards meeting the Committee's request as best they could given the data available and was therefore considered acceptable.	4.17
	The Committee considered the extrapolation of health-state data beyond the trial for the observation arm. The Committee accepted that spontaneous resolution occurs in BRVO and therefore it was not appropriate to include the first 3 months of data for this subgroup. The Committee concluded that 0- to 3-month data for the observation arm of the GENEVA trials for people with CRVO would not have a substantial impact on the overall ICER and therefore accepted the manufacturer's use of 0- to 3-month data in the revised base case.	4.18
	The Committee considered the sensitivity analyses around re-treatment rates provided by the manufacturer in response to requests made by the Committee at the first meeting. The Committee accepted the assumption of mid-point re-treatment rates.	4.19
	The Committee considered the assumptions in the manufacturer's cost-minimisation analysis between dexamethasone and bevacizumab. The Committee concluded that the cost of bevacizumab may have been overestimated, but that there was uncertainty around the number of injections of bevacizumab. The Committee concluded that the assumptions used in the manufacturer's cost-minimisation analysis between dexamethasone and bevacizumab may overestimate the cost of bevacizumab, but the	4.22

	cost may not be as low as suggested by the ERG.	
Incorporation of health-related quality-of-life benefits and utility values Have any potential significant and substantial health-related benefits been identified that were not included in the economic model, and how have they been considered?	The data inputs for the manufacturer's model included utility values estimated using the Visual Function Questionnaire Utility Index (VFQ-UI) and mapped onto the health states using an algorithm from a study eliciting preferences from the general population.	3.9
Are there specific groups of people for whom the technology is particularly cost effective?	The Committee recognised that for a subgroup of people with BRVO with a lesser extent of macular haemorrhage, laser photocoagulation would be a treatment option and that this subgroup is within the licensed indication for dexamethasone. The Committee concluded that because no evidence had been provided for this subgroup, any recommendation for dexamethasone would be restricted to those subgroups of people with BRVO who cannot receive or have already tried and not benefited from laser photocoagulation.	4.6
What are the key drivers of cost effectiveness?	The cost effectiveness of dexamethasone compared with best supportive care was affected by assumptions around re-treatment. The Committee considered that its discussion on the relative cost effectiveness of dexamethasone compared with bevacizumab should focus on the assumptions within the manufacturer's cost-minimisation analysis.	4.15 4.21
Most likely cost-effectiveness estimate (given as an ICER)	The Committee concluded that the decision regarding the cost effectiveness of dexamethasone compared with best supportive care should be based on the manufacturer's revised ICER of £26,300 per QALY gained (including an incremental costs of £5937 and incremental QALYs of 0.23) for all people with RVO. The Committee further concluded that this represented an acceptable level of cost effectiveness in this case and that dexamethasone intravitreal implant for the treatment of RVO represents a cost-effective use of NHS resources when compared with best supportive care.	4.20
Additional factors taken into account		
Patient access schemes (PPRS)	None	
End-of-life	None	

considerations		
Equalities considerations and social value judgements	The Committee concluded that there were no equality issues relating to this appraisal that required addressing in the guidance.	4.25

5 Implementation

- 5.1 The Secretary of State and the Welsh Assembly Minister for Health and Social Services have issued directions to the NHS in England and Wales on implementing NICE technology appraisal guidance. When a NICE technology appraisal recommends use of a drug or treatment, or other technology, the NHS must usually provide funding and resources for it within 3 months of the guidance being published. If the Department of Health issues a variation to the 3-month funding direction, details will be available on the NICE website. When there is no NICE technology appraisal guidance on a drug, treatment or other technology, decisions on funding should be made locally.
- 5.2 NICE has developed tools to help organisations put this guidance into practice (listed below). These are available on our website (www.nice.org.uk/guidance/TA229).
- Costing template and report to estimate the national and local savings and costs associated with implementation.

6 Related NICE guidance

Published

- Ranibizumab and pegaptanib for the treatment of age-related macular degeneration. NICE technology appraisal guidance 155 (2008). Available from www.nice.org.uk/guidance/TA155
- Guidance on the use of photodynamic therapy for age-related macular degeneration. NICE technology appraisal guidance 68 (2003). Available from www.nice.org.uk/guidance/TA68

□nder development

NICE is developing the following guidance (details available from www.nice.org.uk):

- Ranibizumab for the treatment of macular oedema caused by retinal vein occlusion. NICE technology appraisal (publication date to be confirmed).
- Ranibizumab for the treatment of diabetic macular oedema. NICE technology appraisal (publication date to be confirmed).
- Pegaptanib sodium for the treatment of diabetic macular oedema. NICE technology appraisal (publication date to be confirmed).

7 Proposed date for review of guidance

- 7.1 NICE proposes that the guidance on this technology is considered for review by the Guidance Executive in 2014. NICE welcomes comment on this proposed date. The Guidance Executive will decide whether the technology should be reviewed based on information gathered by NICE, and in consultation with consultees and commentators.

Andrew Dillon
Chief Executive
July 2011

Appendix A: Appraisal Committee members, and NICE project team

A *Appraisal Committee members*

The Appraisal Committees are standing advisory committees of NICE. Members are appointed for a 3-year term. A list of the Committee members who took part in the discussions for this appraisal appears below. There are four Appraisal Committees, each with a chair and vice chair. Each Appraisal Committee meets once a month, except in December when there are no meetings. Each Committee considers its own list of technologies, and ongoing topics are not moved between Committees.

Committee members are asked to declare any interests in the technology to be appraised. If it is considered there is a conflict of interest, the member is excluded from participating further in that appraisal.

The minutes of each Appraisal Committee meeting, which include the names of the members who attended and their declarations of interests, are posted on the NICE website.

Dr Kathryn Abel

Reader and Consultant Psychiatrist; and Director of Centre for Women's Mental Health, University of Manchester

Dr David Black

Director of Public Health, Derbyshire County Primary Care Trust

Dr Daniele Bryden

Consultant in Intensive Care Medicine and Anaesthesia, Sheffield Teaching Hospitals NHS Trust

Dr Andrew Burnett

Director for Health Improvement and Medical Director, NHS Barnet, London

David Chandler

Lay member

Dr Mary Cooke

Lecturer, School of Nursing, Midwifery and Social Work, University of Manchester

Dr Chris Cooper

GP, London

Professor Peter Crome

Consultant Geriatrician and Professor of Geriatric Medicine, Keele University

Dr Christine Davey

Research Adviser, North and East Yorkshire Alliance Research and Development Unit, York

Richard Devereaux-Phillips

Public Affairs and Reimbursement Manager UK and Ireland, Medtronic

Professor Rachel A Elliott

Lord Trent Professor of Medicines and Health, University of Nottingham

Dr Wasim Hanif

Consultant Physician and Honorary Senior Lecturer, University Hospital Birmingham

Dr Alan Haycox

Reader in Health Economics, University of Liverpool Management School

Professor Cathy Jackson

Professor of Primary Care Medicine, University of St Andrews

Dr Peter Jackson

Clinical Pharmacologist, University of Sheffield

Henry Marsh

Consultant Neurosurgeon, St George's Hospital, London

Professor Gary McVeigh

Professor of Cardiovascular Medicine, Queens University Belfast and Consultant Physician, Belfast City Hospital

Dr Eugene Milne

Deputy Regional Director of Public Health, North East Strategic Health Authority, Newcastle upon Tyne

Dr Neil Myers

GP, Glasgow

Dr Richard Nakielny

Consultant Radiologist, Sheffield Teaching Hospitals Foundation Trust

Dr Danielle Preedy

Lay member

Dr Martin Price

Head of Outcomes Research, Janssen-Cilag

Ellen Rule

Programme Director, NHS Bristol

Dr Peter Selby

Consultant Physician, Central Manchester University Hospitals NHS Foundation Trust

Dr Surinder Sethi

Consultant in Public Health Medicine, North West Specialised Services Commissioning Team, Warrington

Professor Andrew Stevens

Chair of Appraisal Committee C; Professor of Public Health, University of Birmingham

Dr John Stevens

Lecturer in Bayesian Statistics in Health Economics, School of Health and Related Research, Sheffield

Dr Matt Stevenson

Technical Director, School of Health and Related Research, University of Sheffield

Professor Paul Trueman

Professor of Health Economics, Brunel University, London

Dr Judith Wardle

Lay member

B ***NICE project team***

Each technology appraisal is assigned to a team consisting of one or more health technology analysts (who act as technical leads for the appraisal), a technical adviser and a project manager.

Jennifer Priaulx

Technical Lead

Joanne Holden

Technical Adviser

Lori Farrar

Project Manager

Appendix B: Sources of evidence considered by the Committee

- A The Evidence Review Group (ERG) report for this appraisal was prepared by Aberdeen Health Technology Assessment Group (AbHTAG):

Shyangdan D, Cummins E, Lois N, et al. Dexamethasone implants in the treatment of macular oedema due to retinal vein occlusion: a single technology appraisal. December 2010

- B The following organisations accepted the invitation to participate in this appraisal as consultees and commentators. They were invited to comment on the draft scope, the ERG report and the appraisal consultation document (ACD). Organisations listed in I were also invited to make written submissions. Organisations listed in II and III had the opportunity to give their expert views. Organisations listed in I, II and III also have the opportunity to appeal against the final appraisal determination.

I Manufacturer/sponsor:

- Allergan Ltd U.K.

II Professional/specialist and patient/carer groups:

- Royal National Institute of Blind People (RNIB)
- Royal College of Nursing
- Royal College of Ophthalmologists
- Royal College of Physicians

III Other consultees:

- Department of Health
- Welsh Assembly Government
- Waltham Forest PCT

IV Commentator organisations (did not provide written evidence and without the right of appeal):

- British National Formulary
- Commissioning Support Appraisals Service

- Department of Health, Social Services and Public Safety for Northern Ireland
- NHS Quality Improvement Scotland
- Bristol Myers Squibb (triamcinolone)
- Roche Products (bevacizumab)
- Aberdeen HTA Group
- National Institute for Health Research Health Technology Assessment Programme

C The following individuals were selected from clinical specialist and patient expert nominations from the non-manufacturer/sponsor consultees and commentators. They gave their expert personal view on Dexamethasone intravitreal implant for the treatment of macular oedema caused by retinal vein occlusion by attending the initial Committee discussion and providing written evidence to the Committee. They are invited to comment on the ACD.

- Ian Pearce, Consultant Ophthalmologist, nominated by The Royal College of Ophthalmologists – clinical specialist
- Lady Sandra Taylor, Lead Ophthalmologist research Nurse Manager, nominated by Royal College of Nurses Ophthalmic Nurse Forum – clinical specialist
- Barbara McLaughlan, Campaigns Manager, nominated by Royal National Institute of Blind people – patient expert
- Carol Read, nominated by Royal National Institute of Blind people – patient expert

D Representatives from the following manufacturer/sponsor attended Committee meetings. They contributed only when asked by the Committee chair to clarify specific issues and comment on factual accuracy.

- Allergan Ltd UK

Section 5.0

Coverage Guidances

Percutaneous Interventions for Cervical Spine Pain

Question: How should the HTAS Coverage Guidance regarding percutaneous interventions for cervical spine pain be applied to the Prioritized List?

Question source: Health Technology Assessment Subcommittee

Issue: HTAS approved a new Coverage Guidance at their July, 2014 meeting. This CG needs final approval by HERC at the January, 2015 meeting. The summary of the coverage guidance is shown below:

HERC COVERAGE GUIDANCE

Therapeutic cervical spinal epidural injections are recommended for coverage for cervical spine pain with radiculopathy of six weeks duration (*weak recommendation*) only when all of the following criteria are met:

- documented neuroforaminal stenosis (without infection or neoplasia)
- radicular pain in a corresponding dermatomal distribution,
- pain is intractable and conservative therapy has failed,
- fluoroscopic guidance or CT guidance is utilized,
- interlaminar approach is utilized,
- no more than two injections without clinically meaningful improvement in pain and function, and
- maximum of three injections in six months.

Epidural steroid injections of the cervical spine are not recommended for coverage (*strong recommendation*) for other types of neck pain or for headache.

Therapeutic cervical intraarticular facet joint injections and therapeutic cervical medial branch blocks are not recommended for coverage for facet joint pain (*strong recommendation*).

Facet joint radiofrequency neurotomy is recommended for coverage (*weak recommendation*) only when all the following criteria are met:

- at least 3 months of moderate to severe pain with functional impairment,
- pain is predominantly axial and not associated with radiculopathy,
- conservative therapy has failed, and
- complete or nearly complete pain relief (80% or greater) following fluoroscopically guided, low-volume local anesthetic blocks of the medial branch nerves, performed on two separate occasions using two commonly-used agents with different anticipated durations of action.

This coverage guidance and HERC staff recommended changes to the Prioritized list were discussed at the August VBBS meeting. HERC staff recommended adding cervical epidural steroid injections to 4 spinal lines and facet joint injections to 3 spinal lines, both with new guidelines. After discussion, the VBBS decision was to continue to exclude these procedures. The major concerns of the VBBS were increase in costs if these procedures were covered, and the weak level of evidence supporting the HTAS recommendation to include these procedures.

Percutaneous Interventions for Cervical Spine Pain

HTAS relied on expert testimony and desire to provide options to surgery in recommending coverage. The epidural steroid injections are currently ancillary, whereas the facet joint injections are excluded. The VBBS requested clarification about the level of evidence needed to move a procedure from the ancillary list to the prioritized list. Based on the presumption that procedures on the ancillary list should only not be covered if evidence of harm or ineffectiveness was found, the VBBS decided to cover epidural steroid injections. However, due to the low level of evidence, the VBBS decided to not cover facet joint injections. Later in the meeting, however, the decision was regarding epidural steroid injections was reopened and the general decision was that HERC staff needed to obtain clarification on the level of evidence required to add something to the Prioritized List when it is on the Ancillary List. The decisions on both procedures were tabled until the November VBBS meeting.

At the August HERC meeting, HERC members clarified the level of evidence required to add a previously ancillary procedure to the prioritized list. The HERC felt that ancillary list procedures were unlikely to have been reviewed in the past and should be subject to the same level of evidence requirements to be added to the prioritized list. In other words, Ancillary procedures needed to have good evidence of effectiveness to be added to the prioritized list.

The HTAS summary of evidence is shown in the table below. The procedures in question had very low to low level of evidence of effectiveness.

Expert input: Dr. David Sibell, OHSU Comprehensive Pain Clinic

I've always found it interesting to see what different authorities view as evidence based medicine, when it comes to low back pain. My understanding is that, for acute radiculopathy, there is good evidence that transforaminal epidural steroid injection reduces pain for 6-8 weeks (though it does not reliably affect rate of surgery, which has many other determinants); and that physical therapy and waiting are also effective.

For chronic low back pain, there are repeated studies that demonstrate prolonged improvement in pain and function in patients diagnosed with painful facet arthropathy. The problem is that there are also some very poor quality studies in this literature, and reviewers often group those studies together, even though they use non-standard techniques. It turns out that, if you do the procedure wrong, it doesn't work well...go figure. How likely is it that this treatment would be considered? It makes the highest degree of difference for patients with chronic low back pain (combined with physical therapy, in almost all cases) of anything we do.

PT & CBT also have a considerable evidence base, for both acute and chronic LBP.

I'm glad that medications are getting another look. The evidence is poor for the most commonly used therapies, so there is a chance to do something better for patients.

Percutaneous Interventions for Cervical Spine Pain

Table 1. Efficacy/Effectiveness

Indication	Population characteristics	Comparator	Strength of Evidence	Conclusions
Cervical Epidural Steroid Injections				
Neck pain with disc herniation and radiculitis	Cervical disc herniation or radiculitis, chronic neck pain > 6 months duration, no spinal stenosis unless accompanied by disc herniation, no prior cervical spine surgery, opioid use stable/ controlled	Placebo epidural injection (local anaesthetic)	Very low	No benefit in terms of pain, function, or opioid use at both three and twelve months or on employment at twelve months based on data from one RCT.
Neck pain without disc herniation and radiculitis	Negative diagnosis of facet joint pain (by use of local anaesthetic blocks), absence of disc herniation/ radiculitis or spinal stenosis, chronic neck/arm pain > 6 months duration, opioid use stable/ controlled	Placebo epidural injection (local anaesthetic)	Very low	No benefit in terms of pain, function, or opioid use at both three and twelve months or on employment at twelve months based on data from one RCT.
Neck pain with disc compression and radiculitis	Chronic resistant cervicobrachialgia, continued pre-study treatments (medications, PT). Full text only available for purchase	Intramuscular injection (saline and local anaesthetic)	Very low	Epidural injections were superior to intramuscular injections in the posterior neck in terms of pain, analgesic use, and employment at one week and twelve months based on data from one RCT.
Post-surgery syndrome	Cervical post-surgery syndrome, surgery > 1 year previously with continued pain for at least 6 months after surgery, opioid use	Placebo epidural injection (local anaesthetic)	Very low	No benefit in terms of pain or disability at 3, 6 and 12 months based on data from one RCT.

Percutaneous Interventions for Cervical Spine Pain

Indication	Population characteristics	Comparator	Strength of Evidence	Conclusions
	stable/ controlled			
Cervical spinal stenosis	Cervical central spinal stenosis with or w/o foraminal stenosis, age > 30, chronic neck pain of at least 6 on a scale of 1-10, at least 6 months duration, failed to improve with conservative management (PT, chiro, exercise, drugs, bed rest), opioid use stable/ controlled	Placebo epidural injection (local anaesthetic)	Very low	No benefit in terms of pain or disability at 3, 6 and 12 months based on data from one RCT.
Cervical Intraarticular Facet Joint Steroid Injections				
Confirmed facet joint pain	Neck pain lasting > 3 months due to MVA with documented positive response to medial branch block	Placebo facet injection (local anaesthetic)	Very low	No benefit in terms of the length of pain relief based on one RCT. No long-term data was reported.
Cervical Medial Branch Blocks (local anaesthetic + steroid, ± Sarapin*)				
Confirmed facet joint pain	Non-specific neck pain > 6 months duration, confirmed facet joint pain (response to local anaesthetic blocks), no radicular symptoms, failed conservative management (PT, chiro, exercise, drugs, bed rest), opioid use not heavy, continued previous interventions	Placebo injection (local anaesthetic ± Sarapin)	Very low	No benefit in terms of pain or function at both three and twelve months or on opioid use or employment at twelve months based on one RCT.
Facet Neurotomy				

Percutaneous Interventions for Cervical Spine Pain

Indication	Population characteristics	Comparator	Strength of Evidence	Conclusions
Confirmed facet joint pain	Failed conservative management (mean duration of pain 34 to 44 months), responded to 3 MBBs (100% pain relief)	Sham RF neurotomy	Very low	Significantly more patients in the RF neurotomy group had freedom from “accustomed pain” (risk difference, 50% (95% CI, 18% to 82%) (P = 0.01)
Cervicogenic headache	Cervicogenic HA of at least 2 years duration, rated at least 50 out of 100 on VAS, considerable pain at least 2 days/wk, no prior neck surgery	Injection of major occipital nerve	Low	No difference between groups in headache relief or a composite measure of success at 2 months
Facet joint pain	Patients with previously successful RF neurotomy	None – 3 case series	Very low	Patients are likely to experience a similar response to subsequent RF neurotomy procedures as they experienced during their first RF neurotomy procedure

Percutaneous Interventions for Cervical Spine Pain

HERC staff recommendations:

Option 1 (following previous VBBS discussion)

- 1) Add 63210 (Injection(s), of diagnostic or therapeutic substance(s) (including anesthetic, antispasmodic, opioid, steroid, other solution), not including neurolytic substances, including needle or catheter placement, includes contrast for localization when performed, epidural or subarachnoid; cervical or thoracic) to the Excluded List
 - a. Advise DMAP to remove 63210 from the Ancillary List
- 2) Keep 64490-64492 (Injection(s), diagnostic or therapeutic agent, paravertebral facet (zygapophyseal) joint (or nerves innervating that joint) with image guidance (flouro or CT), cervical) and 64633 and 64634 (Destruction by neurolytic agent, paravertebral facet joint nerve(s), with imaging guidance (fluoroscopy or CT); cervical or thoracic, single or additional facet joints) on the Excluded List
- 3) Add an entry to the new "excluded list" for CPT 63210, 64633-64634, 64479-64480, 64490-64495

Option 2 (following HTAS recommendations with August 2014 VBBS guideline edits)

- 1) Add 63210 (Injection(s), of diagnostic or therapeutic substance(s) (including anesthetic, antispasmodic, opioid, steroid, other solution), not including neurolytic substances, including needle or catheter placement, includes contrast for localization when performed, epidural or subarachnoid; cervical or thoracic) to lines 374 DISORDERS OF SPINE WITH NEUROLOGIC IMPAIRMENT, 412 SPINAL DEFORMITY, CLINICALLY SIGNIFICANT, 545 ACUTE AND CHRONIC DISORDERS OF SPINE WITHOUT NEUROLOGIC IMPAIRMENT, 588 SPINAL DEFORMITY, NOT CLINICALLY SIGNIFICANT
 - a. Add a new guideline note regarding cervical epidural steroid injections as shown below
 - b. Advise DMAP to remove 63210 from the Ancillary List
- 2) Add an entry to the new "excluded list" for CPT 64479-64480, 0213T– 0218T
- 3) Add 64490-64492 (Injection(s), diagnostic or therapeutic agent, paravertebral facet (zygapophyseal) joint (or nerves innervating that joint) with image guidance (flouro or CT), cervical) and 64633 and 64634 (Destruction by neurolytic agent, paravertebral facet joint nerve(s), with imaging guidance (fluoroscopy or CT); cervical or thoracic, single or additional facet joints) to lines 412, 545, 588
 - a. Advise DMAP to remove 64490-64492, 64633 and 64634 from the Excluded List
 - a. Adopt a new guideline note regarding facet joint radiofrequency neurotomy as shown below

GUIDELINE NOTE XXX CERVICAL SPINAL EPIDURAL INJECTIONS

Lines 374,412,545,588

Cervical spinal epidural injections (CPT 62310) are **not** included on these lines for use in diagnostic testing.

Cervical spinal epidural injections (CPT 62310) for therapeutic purposes are included on these for cervical spine pain with radiculopathy of 6 weeks duration when all of the following criteria are met:

- 1) documented neuroforaminal stenosis (without infection or neoplasia)
- 2) radicular pain in a corresponding dermatomal distribution,
- 3) pain is intractable and conservative therapy has failed,

Percutaneous Interventions for Cervical Spine Pain

- 4) fluoroscopic guidance or CT guidance is utilized,
- 5) interlaminar approach is utilized,
- 6) no more than two injections without clinically meaningful improvement in pain and function, AND
- 7) maximum of three injections in six months.

Epidural steroid injections of the cervical spine are not included on these lines for treatment of other types of neck pain or for headache.

GUIDELINE NOTE XXX CERVICAL FACET JOINT RADIOFREQUENCY NEUROTOMY AND DIAGNOSTIC INJECTIONS

Lines 412,545,588

Cervical medial branch block injections (CPT 64490-64492) are included on these lines only when done as a diagnostic precursor to facet joint radiofrequency neurotomy as specified in #4 below.

Cervical facet joint radiofrequency neurotomy (CPT 64633, 64634) is included on these lines only when all the following criteria are met:

- 1) at least 3 months of moderate to severe pain with functional impairment,
- 2) pain is predominantly axial and not associated with radiculopathy,
- 3) conservative therapy has failed, and
- 4) complete or nearly complete pain relief (80% or greater) following fluoroscopically guided, low-volume local anesthetic blocks of the medial branch nerves, performed on two separate occasions using two commonly-used agents with different anticipated durations of action.

Thoracic, lumbar and sacral paravertebral facet joint diagnostic injections and radiofrequency neurotomy are not included on these lines.

Percutaneous Interventions for Cervical Spine Pain

Percutaneous Interventions for Cervical Spine Pain

Appendix—materials previously provided to the VBBS

Current Prioritized List status (note: line numbers refer to January 1, 2015 Prioritized List)

CODES	DESCRIPTION	LOCATION	STAFF RECOMMENDATION
ICD-9 Diagnosis Codes			
720.9	Unspecified inflammatory spondylopathy	50 RHEUMATOID ARTHRITIS AND OTHER INFLAMMATORY POLYARTHROPATHIES	No change
721	Spondylosis , various	412 SPINAL DEFORMITY, CLINICALLY SIGNIFICANT OR 545 ACUTE AND CHRONIC DISORDERS OF SPINE WITHOUT NEUROLOGIC IMPAIRMENT	No change
722	Degeneration/displacement intervertebral disc, various	374 DISORDERS OF SPINE WITH NEUROLOGIC IMPAIRMENT 545	No change
723	Spinal stenosis, brachial neuritis or radiculitis, various	412 515 PERIPHERAL NERVE DISORDERS Therapy: MEDICAL THERAPY 541 PERIPHERAL NERVE DISORDERS Therapy: SURGICAL TREATMENT 545 588 SPINAL DEFORMITY, NOT CLINICALLY SIGNIFICANT	No change
724	Spinal stenosis, various	374,412,545,588	No change
738.4	Acquired spondylolisthesis	588	No change
ICD-10 Diagnosis Codes			
M43.00-3	Spondylolysis	588	No change
M43.10-3	Spondylolisthesis	588	No change
M47.10-3	Other spondylosis with myelopathy	412	No change
M47.20-3	Other spondylosis with radiculopathy	374	No change
M47.811-3, M47.819	Spondylosis without myelopathy or radiculopathy	545	No change
M47.891-3	Other spondylosis	545	No change
M48.00-3	Spinal stenosis	412,588	No change
M50	Cervical disc disorders	374	No change

Percutaneous Interventions for Cervical Spine Pain

M54.10-3	Radiculopathy	374	No change
M54.2	Cervicalgia	374,545	No change
CPT Codes			
62310	Injection(s), of diagnostic or therapeutic substance(s) (including anesthetic, antispasmodic, opioid, steroid, other solution), not including neurolytic substances, including needle or catheter placement, includes contrast for localization when performed, epidural or subarachnoid; cervical or thoracic	Ancillary	Add to lines 374,412, 545,588 with a new guideline note Advise DMAP to remove from the Ancillary List
64479-80	Injection(s), anesthetic agent and/or steroid, transforaminal epidural, with imaging guidance; cervical or thoracic, single level and each additional	Excluded	Keep Excluded Add entry to new excluded list
64490-64492	Injection(s), diagnostic or therapeutic agent, paravertebral facet (zygapophyseal) joint (or nerves innervating that joint) with image guidance (fluoro or CT), cervical, thoracic, lumbar or sacral; single, second or third level	Excluded	Add 64490-64492 to lines 412,545,588 with a new guideline note Advise DMAP to remove 64490-64492 from the Excluded List
64633-64634	Destruction by neurolytic agent, paravertebral facet (zygapophyseal) joint (or nerves innervating that joint) with image guidance (fluoro or CT), cervical, thoracic, lumbar or sacral; single, second or third level	Excluded	Add 64633 and 64634 to lines 412, 545,588 with a new guideline note Advise DMAP to remove 64633 and 64634 from the Excluded List
0213T–0215T	Injection(s), diagnostic or therapeutic agent, paravertebral facet (zygapophyseal) joint (or nerves innervating that joint) with ultrasound guidance, cervical, thoracic, lumbar or sacral; single, second or third level	Excluded	No change
77003	Fluoroguide for spine injection	Ancillary	No change

Percutaneous Interventions for Cervical Spine Pain

Relevant Guidelines

GUIDELINE NOTE 37, DISORDERS OF SPINE WITH NEUROLOGIC IMPAIRMENT

Lines 374,545

Diagnoses are included on Line 374 when neurologic impairment or radiculopathy is present, as defined as:

- A) Markedly abnormal reflexes
- B) Segmental muscle weakness
- C) Segmental sensory loss
- D) EMG or NCV evidence of nerve root impingement
- E) Cauda equina syndrome,
- F) Neurogenic bowel or bladder
- G) Long tract abnormalities

Otherwise, disorders of spine not meeting these criteria (e.g. pain alone) fall on Line 545.

DIAGNOSTIC GUIDELINE D4, ADVANCED IMAGING FOR LOW BACK PAIN

In patients with non-specific low back pain and no “red flag” conditions [see Table D4], imaging is not a covered service; otherwise work up is covered as shown in the table.

Electromyography (CPT 96002-4) is not covered for non-specific low back pain.

Table D4

Low Back Pain - Potentially Serious Conditions (“Red Flags”) and Recommendations for Initial Diagnostic Work-up

Possible cause	Key features on history or physical examination	Imaging*	Additional studies*
Cancer	History of cancer with new onset of LBP	MRI	ESR
	- Unexplained weight loss - Failure to improve after 1 month - Age >50 years - Symptoms such as painless neurologic deficit, night pain or pain increased in supine position	Lumbosacral plain radiography	
	Multiple risk factors for cancer present	Plain radiography or MRI	
Spinal column infection	- Fever - Intravenous drug use - Recent infection	MRI	ESR and/or CRP
Cauda equina syndrome	- Urinary retention - Motor deficits at multiple levels - Fecal incontinence - Saddle anesthesia	MRI	None
Vertebral compression fracture	- History of osteoporosis - Use of corticosteroids - Older age	Lumbosacral plain radiography	None

Percutaneous Interventions for Cervical Spine Pain

Possible cause	Key features on history or physical examination	Imaging*	Additional studies*
Ankylosing spondylitis	• Morning stiffness • Improvement with exercise • Alternating buttock pain • Awakening due to back pain during the second part of the night • Younger age	Anterior-posterior pelvis plain radiography	ESR and/or CRP, HLA-B27
Nerve compression/ disorders (e.g. herniated disc with radiculopathy)	• Back pain with leg pain in an L4, L5, or S1 nerve root distribution present < 1 month • Positive straight-leg-raise test or crossed straight-leg-raise test	None	None
	• Radiculopathic signs** present >1 month • Severe/progressive neurologic deficits (such as foot drop), progressive motor weakness	MRI***	Consider EMG/NCV
Spinal stenosis	• Radiating leg pain • Older age • Pain usually relieved with sitting (Pseudoclaudication a weak predictor)	None	None
	• Spinal stenosis symptoms present >1 month	MRI**	Consider EMG/NCV

* Level of evidence for diagnostic evaluation is variable

** Radiculopathic signs are defined for the purposes of this guideline as in Guideline Note 37 with any of the following:

- A. Markedly abnormal reflexes
- B. Segmental muscle weakness
- C. Segmental sensory loss
- D. EMG or NCV evidence of nerve root impingement
- E. Cauda equina syndrome,
- F. Neurogenic bowel or bladder
- G. Long tract abnormalities

*** Only if patient is a potential candidate for surgery or epidural steroid injection

Red Flag: Red flags are findings from the history and physical examination that may be associated with a higher risk of serious disorders. CRP = C-reactive protein; EMG = electromyography; ESR = erythrocyte sedimentation rate; MRI = magnetic resonance imaging; NCV = nerve conduction velocity.

Extracted and modified from Chou R, Qaseem A, Snow V, et al: *Diagnosis and Treatment of Low Back Pain: A Joint Clinical Practice Guideline from the American College of Physicians and the American Pain Society*. Ann Intern Med. 2007; 147:478-491.

Utilization data:

Based on claims data analysis for the past 6 months, 63210 paired with diagnoses from the following lines (January 1, 2015 PL):

154 CERVICAL VERTEBRAL DISLOCATIONS/FRACTURES, OPEN OR CLOSED; OTHER VERTEBRAL DISLOCATIONS/FRACTURES, OPEN OR UNSTABLE; SPINAL CORD INJURIES WITH OR WITHOUT EVIDENCE OF VERTEBRAL INJURY

374 DISORDERS OF SPINE WITH NEUROLOGIC IMPAIRMENT

412 SPINAL DEFORMITY, CLINICALLY SIGNIFICANT

484 CLOSED DISLOCATIONS/FRACTURES OF NON-CERVICAL VERTEBRAL COLUMN WITHOUT NEUROLOGIC INJURY OR STRUCTURAL INSTABILITY

515 PERIPHERAL NERVE DISORDERS

545 ACUTE AND CHRONIC DISORDERS OF SPINE WITHOUT NEUROLOGIC IMPAIRMENT

588 SPINAL DEFORMITY, NOT CLINICALLY SIGNIFICANT

612 DISORDERS OF SOFT TISSUE

Percutaneous Interventions for Cervical Spine Pain

HERC staff recommended guidelines presented at the August VBBS meeting with meeting edits

GUIDELINE NOTE XXX CERVICAL SPINAL EPIDURAL INJECTIONS

Lines 374,412,545,588

Cervical spinal epidural injections (CPT 62310) are **not** included on these lines for use in diagnostic testing.

Cervical spinal epidural injections (CPT 62310) for therapeutic purposes are included on these [lines](#) for cervical spine pain with radiculopathy of 6 weeks duration when all of the following criteria are met:

- 8) documented neuroforaminal stenosis (without infection or neoplasia)
- 9) radicular pain in a corresponding dermatomal distribution,
- 10) pain is intractable and conservative therapy has failed,
- 11) fluoroscopic guidance or CT guidance is utilized,
- 12) interlaminar approach is utilized,
- 13) no more than two injections without clinically meaningful improvement in pain and function, AND
- 14) maximum of three injections in six months.

Epidural steroid injections of the cervical spine are not included on these lines for treatment of other types of neck pain or for headache.

GUIDELINE NOTE XXX [Cervical](#) FACET JOINT RADIOFREQUENCY NEUROTOMY AND DIAGNOSTIC INJECTIONS

Lines 412,545,588

[Cervical medial branch block](#) ~~Paravertebral facet joint diagnostic~~ injections (CPT 64490-64492) are included on these lines only when done as a diagnostic precursor to facet joint radiofrequency neurotomy as specified in #4 below.

[Cervical](#) Facet joint radiofrequency neurotomy (CPT 64633, 64634) is included on these lines only when all the following criteria are met:

- 1) at least 3 months of moderate to severe pain with functional impairment,
- 2) pain is predominantly axial and not associated with radiculopathy,
- 3) conservative therapy has failed, and
- 4) complete or nearly complete pain relief (80% or greater) following fluoroscopically guided, low-volume local anesthetic blocks of the medial branch nerves, performed on two separate occasions using two commonly-used agents with different anticipated durations of action.

[Thoracic, lumbar and sacral paravertebral facet joint diagnostic injections and radiofrequency neurotomy are not included on these lines.](#)

HERC Coverage Guidance – Percutaneous Interventions for Cervical Spine Pain

Disposition of Public Comments

Table of Contents

Commenters.....	1
Public Comments	2
References Provided by Commenters	4

Commenters

Identification	Stakeholder
A	North American Spine Society (NASS), Burr Ridge, IL <i>[Submitted June 6, 2014]</i>

HERC Coverage Guidance – Percutaneous Interventions for Cervical Spine Pain

Disposition of Public Comments

Public Comments

Ident.	#	Comment	Disposition
A	1	NASS agrees with a majority of the HERC guidance regarding coverage for therapeutic cervical spinal epidural injections for chronic cervical pain with radiculopathy. Regarding indications, in addition to the documentation of herniated intervertebral disc, other anatomic causes of radiculopathy should also be included such as: • Neuroforaminal stenosis; • Central stenosis, disc protrusions; and/or • Segmental spondylosis with radicular pain. NASS strongly agrees that all cervical epidural injections or selective nerve root blocks should be performed with fluoroscopic or CT-guidance.	Thank you for taking the time to comment. HTAS acknowledges that there are other anatomic causes of radiculopathy and has changed the criteria in the coverage recommendation to the following: “documented neuroforaminal stenosis (without infection or neoplasia)”
	2	NASS would like to clarify also that an epidural injection can be appropriately performed by either an interlaminar approach or a transforaminal approach. The HERC reference to a translaminar approach is incorrect as there is no translaminar approach with these injections and the terminology to reference either an interlaminar or a transforaminal approach should be corrected in your policy guidance. NASS also wishes to clarify, as referenced in your policy, the utility of diagnostic spinal nerve (root) blocks in the treatment and/or diagnosis of cervical radicular pain. These blocks use a foraminal approach and share the same CPT codes (64479-64480) as the technically similar transforaminal epidural injection.	These terms translaminar and interlaminar are used interchangeably in the WA HTA report and in CMS coverage decisions. Terminology changed to reflect your suggestions in the body of the guidance; however, when this term is used in coverage policies, no change has been made. The diagnostic blocks referred to in the coverage guidance are medial branch blocks, and the CPT codes utilized (64490-2) are different than those referred to by the commenter. Diagnostic spinal nerve root blocks (64479-80) were not reviewed in this document. Text in the guidance document has been clarified and

HERC Coverage Guidance – Percutaneous Interventions for Cervical Spine Pain

Disposition of Public Comments

Ident.	#	Comment	Disposition
			CPT codes added.
	3	Although many patients may respond to conservative treatment (rest, NSAIDs, PT), in patients with moderate to severe radicular pain with resulting functional limitations (i.e., inability to work, etc.) and appropriate correlated imaging findings, cervical epidural injections are indicated in the acute and sub-acute phases of treatment and is an appropriate conservative treatment option in this patient population.	All studies in the evidence review included only patients with chronic neck pain. When this was defined, it was pain of > 6 months duration or more. The background section of this document reports the definition of chronic to be greater than 3 months. HTAS has revised the box language to specify that coverage is recommended for pain of at least 6 weeks duration.
	4	NASS would like to recommend the Health Evidence Review Commission to include a transparent evidentiary table that will allow users to access and review available supporting literature as well as provide a definition of chronic for further clarification.	Links to the supporting literature, when publically available, are provided on page 2 of the guidance document. See comment 3 regarding definition of chronic.
	5	Please click on the link below for the NASS coverage recommendations on cervical epidural injections and diagnostic spinal nerve blocks for your reference: https://www.spine.org/Pages/PolicyPractice/Coverage/CoverageRecommendations.aspx	Thank you for providing this information.

HERC Coverage Guidance – Percutaneous Interventions for Cervical Spine Pain Disposition of Public Comments

References Provided by Commenters

Commenter	References
A	NASS. (2014). Cervical Epidural Injections And Diagnostic Spinal Nerve Block. https://www.spine.org/Documents/PolicyPractice/CoverageRecommendations/CervicalEpiduralInjections.pdf

HEALTH EVIDENCE REVIEW COMMISSION (HERC)

COVERAGE GUIDANCE: PERCUTANEOUS INTERVENTIONS FOR CERVICAL SPINE PAIN

DRAFT for 8/14/2014 VbBS Meeting Materials

HERC COVERAGE GUIDANCE

Therapeutic cervical spinal epidural injections are recommended for coverage for cervical spine pain with radiculopathy of six weeks duration (*weak recommendation*) only when all of the following criteria are met:

- documented neuroforaminal stenosis (without infection or neoplasia)
- radicular pain in a corresponding dermatomal distribution,
- pain is intractable and conservative therapy has failed,
- fluoroscopic guidance or CT guidance is utilized,
- interlaminar approach is utilized,
- no more than two injections without clinically meaningful improvement in pain and function, and
- maximum of three injections in six months.

Epidural steroid injections of the cervical spine are not recommended for coverage (*strong recommendation*) for other types of neck pain or for headache.

Therapeutic cervical intraarticular facet joint injections and therapeutic cervical medial branch blocks are not recommended for coverage for facet joint pain (*strong recommendation*).

Facet joint radiofrequency neurotomy is recommended for coverage (*weak recommendation*) only when all the following criteria are met:

- at least 3 months of moderate to severe pain with functional impairment,
- pain is predominantly axial and not associated with radiculopathy,
- conservative therapy has failed, and
- complete or nearly complete pain relief (80% or greater) following fluoroscopically guided, low-volume local anesthetic blocks of the medial branch nerves, performed on two separate occasions using two commonly-used agents with different anticipated durations of action.

Note: Definitions for strength of recommendation are provided in Appendix A GRADE Element Description

RATIONALE FOR GUIDANCE DEVELOPMENT

The HERC selects topics for guideline development or technology assessment based on the following principles:

- Represents a significant burden of disease
- Represents important uncertainty with regard to efficacy or harms
- Represents important variation or controversy in clinical care
- Represents high costs, significant economic impact
- Topic is of high public interest

Coverage guidance development follows to translate the evidence review to a policy decision. Coverage guidance may be based on an evidence-based guideline developed by the Evidence-based Guideline Subcommittee or a health technology assessment developed by the Health Technology Assessment Subcommittee. In addition, coverage guidance may utilize an existing evidence report produced by one of HERC's trusted sources, generally within the last three years.

EVIDENCE SOURCES

Trusted Sources

Hashimoto, R., Raich, A., Ecker, E., Henrikson, N., Wallace, L., Dettori, J., & Chou, R. (2011). *Spinal injections*. Olympia, WA: Washington Health Technology Assessment Program. Retrieved from http://www.hta.hca.wa.gov/spinal_injections.html

Little, A., Pettinari, C., Vandegriff, S., Leof, A., Rahman, B., Zoller, E., Kriz, H., Gerrity, M., & King, V. (2013). *Spinal injections: Update to the March 2011 WA HTA report*. Portland, OR: Center for Evidence-based Policy, Oregon Health & Science University.

Hashimoto, R., Holmer, H., Sherry, N., & Skelly, A. (2014). *Facet neurotomy*. Olympia, WA: Washington Health Technology Assessment Program. Retrieved from <http://www.hca.wa.gov/hta/Pages/neurotomy.aspx>

Additional Sources

Cohen, S.P., Bicket, M.C., Jamison, D., Wilkinson, I., & Rathmell, J.P. (2013). Epidural steroids: A comprehensive, evidence-based review. *Regional Anesthesia and Pain Medicine*, 38(3), 175-200. doi: 10.1097/AAP.0b013e31828ea086.

Lee, S.H., Kim, K.T., Kim, D.H., Lee, B.J., Son, E.S., & Kwack, Y.H. (2012). Clinical outcomes of cervical radiculopathy following epidural steroid injection: a prospective study with follow-up for more than 2 years. *Spine (Phila Pa 1976)*, 37(12), 1041-1047. doi: 10.1097/BRS.0b013e31823b4d1f.

The summary of evidence in this document is derived directly from these evidence sources, and portions are extracted verbatim.

SUMMARY OF EVIDENCE

Clinical Background

Chronic neck pain (defined as neck pain that persists more than three months) is common. Risk factors for spinal pain, including neck pain, include increasing age, spinal or disc degeneration, poor posture, anxiety, depression, and injuries. It can be attributed to a number of pathologic changes in the spine, including the following:

- Degenerative disc disease
- Herniated nucleus pulposus
- Spinal stenosis
- Facet joint syndrome
- Whiplash

Medical treatments available for chronic neck pain include non-invasive interventions such as physical therapy, pharmacologic therapy, psychological therapy, exercise, and spinal manipulation. For those patients with inadequate response to those treatments, spinal injections may be considered. Spinal injections involve the injection of an anti-inflammatory agent such as a steroid and/or an anesthetic agent into the spaces in and around the spinal nerves and joints. Types of spinal injection include epidural and facet joint injections and medial branch blocks. The latter two are generally used for pain believed to originate in the facet joint. It has been estimated that the point prevalences of facet joint pain are 10-15% in the low back and 45-55% in the neck. The primary symptom suggestive of facet joint pain is paraspinal tenderness at the affected facet joints, although other symptoms may be present as well. Diagnosis of facet joint pain cannot be accurately made by physical exam or imaging studies alone and diagnostic nerve blocks (medial branch blocks) may be the most accurate assessment method. Epidural injections deliver medication into the epidural space of the spine to decrease inflammation of the nerve root. Two approaches are possible in the neck. The interlaminar approach involves placement of the needle between the lamina of the vertebrae, delivering medication to both the right and left sides of the inflamed area. The transforaminal approach involves placement of the needle in the neural foramen, treating one side at a time.

Facet joint injections deliver the medications (anesthetic with or without a corticosteroid) into the facet joints and include several approaches. Medial branch blocks involve injection of the medication into the area of the medial branch of the posterior primary

ramus. Prior to steroid injections, controlled diagnostic blocks (medial branch blocks) of the joint or the nerves that supply the joint are often performed using local anesthetic. A positive block indicates that pain is eliminated and the affected nerve has been identified as the source of pain.

Once the facet joint is determined to be the source of pain as indicated by a positive diagnostic block, then prolonged pain relief may be achieved with destruction of the nerves to the affected joint in a procedure called facet neurotomy. Neurotomy does not cure the source of pain, but instead cuts off the pain signal from the brain by damaging the nerve. Different types of facet neurotomy are available, but the most common type employs radiofrequency needles to destroy the nerve tissue with heat generated by an electric current. During this procedure, the skin is anesthetized with a local anesthetic and the radiofrequency needles are advanced using guidance to confirm that the needles are properly positioned at the affected nerves. Then a radiofrequency current is applied to disrupt the ability of the nerves to transmit pain signals to the brain.

There are three types of radiofrequency (RF) neurotomy: standard, pulsed and cooled. Pulsed RF neurotomy delivers short bursts of radiofrequency current rather than the continuous flow utilized in non-pulsed RF neurotomy. Pulsed neurotomy allows the nerve tissue to cool between bursts, and is reported to reduce the destruction of neighboring tissue. Some other names used for this procedure include percutaneous radiofrequency denervation, nerve ablation, neurolysis, medial branch neurotomy, medial branch rhizotomy, and articular rhizolysis. Other types of facet neurotomy involve chemical ablation (application of ethyl alcohol, phenol, or sodium morrhuate; cryoablation (application of extreme cold); or laser ablation (application of laser beams) of the medial branch nerves to destroy the nerves and reduce or eliminate pain.

Evidence Review

Trusted Sources

These reports include studies of adults with sub-acute or chronic neck pain due to conditions including (but not limited to) degenerative disc disease, radiculopathy, disc herniation, spinal stenosis and facet joint pain. Evaluated treatments included epidural injections, intraarticular facet injections, medial branch blocks and facet neurotomy. Studies reporting on diagnostic injections (selective nerve root blocks), extraspinal injections (except as a comparator), chemonucleolysis, intradiscal electrothermal therapy and coblation nucleoplasty were excluded. Comparators included any placebo injection (water, saline, local anaesthetic) or non-placebo controls. For efficacy, only randomized controlled trials (RCTs) were included.

A total of seven RCTs were identified that reported on efficacy. Studies were limited to the following populations:

- Chronic neck pain without disc herniation or radiculitis (one study)
- Chronic neck pain with disc herniation or radiculitis (one study)
- Chronic neck pain with resistant cervicobrachialgia (one study)
- Chronic neck pain of facet origin (two studies)
- Cervical post-surgery syndrome (one study)
- Cervical spinal stenosis (one study)

Most of the included studies evaluated one injection technique (interlaminar epidural, facet joint or medial branch block) and compared steroid plus local anaesthetic injection to local anaesthetic alone. However, the trial of patients with resistant cervicobrachialgia compared epidural steroid plus local anaesthetic injection to posterior neck muscle injection with the same substances. The one trial that evaluated facet joint injections compared steroid injection (without local anaesthetic) to local anaesthetic injection.

Efficacy/Effectiveness

For trials comparing cervical epidural steroid injections with:

- Placebo injections for neck pain with disc herniation and radiculitis, there is no benefit based on data from one lower-quality RCT (strength of evidence = very low).
- Placebo injections for neck pain without disc herniation and radiculitis, there is no benefit based on data from one lower-quality RCT (strength of evidence = very low).
- Intramuscular injections for neck pain with disc compression and radiculitis, there is evidence that epidural injections were superior based on data from one lower-quality RCT (strength of evidence = very low).
- Placebo injections for post-surgery syndrome, there is no benefit based on data from one fair quality RCT (strength of evidence = very low).
- Placebo injections for cervical spinal stenosis, there is no benefit based on data from one fair quality RCT (strength of evidence = very low).

For trials comparing cervical intraarticular facet joint steroid injections with:

- Placebo injections for confirmed facet joint pain, there is no benefit based on data from one lower-quality RCT (strength of evidence = very low).

For trials comparing cervical medial branch blocks with:

- Placebo injections for confirmed facet joint pain, there is no benefit based on data from one lower-quality RCT (strength of evidence = very low).

For trials comparing different approaches for administering cervical epidural steroids:

- There were no significant differences between anterolateral and posterolateral approaches in estimated change from baseline in pain score or disability index at two weeks or two months. There is no apparent difference in efficacy based on differing approaches for administering cervical epidural steroids (strength of evidence = very low).

For trials comparing RF neurotomy with:

- Sham neurotomy in patients with confirmed facet joint pain (based on 100% response to three medial branch blocks), significantly more patients in the RF neurotomy group had freedom from “accustomed pain” (risk difference, 50% (95% CI, 18% to 82%) (P = 0.0110) based on evidence from one small (N=24) RCT (strength of evidence = very low).

Table 1. Efficacy/Effectiveness

Indication	Population characteristics	Comparator	Strength of Evidence	Conclusions
Cervical Epidural Steroid Injections				
Neck pain with disc herniation and radiculitis	Cervical disc herniation or radiculitis, chronic neck pain > 6 months duration, no spinal stenosis unless accompanied by disc herniation, no prior cervical spine surgery, opioid use stable/ controlled	Placebo epidural injection (local anaesthetic)	Very low	No benefit in terms of pain, function, or opioid use at both three and twelve months or on employment at twelve months based on data from one RCT.
Neck pain without disc herniation and radiculitis	Negative diagnosis of facet joint pain (by use of local anaesthetic blocks), absence of disc herniation/ radiculitis or spinal stenosis, chronic neck/arm pain > 6 months duration, opioid use stable/ controlled	Placebo epidural injection (local anaesthetic)	Very low	No benefit in terms of pain, function, or opioid use at both three and twelve months or on employment at twelve months based on data from one RCT.

Indication	Population characteristics	Comparator	Strength of Evidence	Conclusions
Neck pain with disc compression and radiculitis	Chronic resistant cervicobrachialgia, continued pre-study treatments (medications, PT). Full text only available for purchase	Intramuscular injection (saline and local anaesthetic)	Very low	Epidural injections were superior to intramuscular injections in the posterior neck in terms of pain, analgesic use, and employment at one week and twelve months based on data from one RCT.
Post-surgery syndrome	Cervical post-surgery syndrome, surgery > 1 year previously with continued pain for at least 6 months after surgery, opioid use stable/ controlled	Placebo epidural injection (local anaesthetic)	Very low	No benefit in terms of pain or disability at 3, 6 and 12 months based on data from one RCT.
Cervical spinal stenosis	Cervical central spinal stenosis with or w/o foraminal stenosis, age > 30, chronic neck pain of at least 6 on a scale of 1-10, at least 6 months duration, failed to improve with conservative management (PT, chiro, exercise, drugs, bed rest), opioid use stable/ controlled	Placebo epidural injection (local anaesthetic)	Very low	No benefit in terms of pain or disability at 3, 6 and 12 months based on data from one RCT.
Cervical Intraarticular Facet Joint Steroid Injections				
Confirmed facet joint pain	Neck pain lasting > 3 months due to MVA with documented positive response to medial branch block	Placebo facet injection (local anaesthetic)	Very low	No benefit in terms of the length of pain relief based on one RCT. No long-term data was reported.
Cervical Medial Branch Blocks (local anaesthetic + steroid, ± Sarapin*)				
Confirmed facet joint	Non-specific neck pain > 6 months	Placebo injection	Very low	No benefit in terms of pain or function at both three

Indication	Population characteristics	Comparator	Strength of Evidence	Conclusions
pain	duration, confirmed facet joint pain (response to local anaesthetic blocks), no radicular symptoms, failed conservative management (PT, chiro, exercise, drugs, bed rest), opioid use not heavy, continued previous interventions	(local anaesthetic ± Sarapin)		and twelve months or on opioid use or employment at twelve months based on one RCT.
Facet Neurotomy				
Confirmed facet joint pain	Failed conservative management (mean duration of pain 34 to 44 months), responded to 3 MBBs (100% pain relief)	Sham RF neurotomy	Very low	Significantly more patients in the RF neurotomy group had freedom from “accustomed pain” (risk difference, 50% (95% CI, 18% to 82%) (P = 0.01)
Cervicogenic headache	Cervicogenic HA of at least 2 years duration, rated at least 50 out of 100 on VAS, considerable pain at least 2 days/wk, no prior neck surgery	Injection of major occipital nerve	Low	No difference between groups in headache relief or a composite measure of success at 2 months
Facet joint pain	Patients with previously successful RF neurotomy	None – 3 case series	Very low	Patients are likely to experience a similar response to subsequent RF neurotomy procedures as they experienced during their first RF neurotomy procedure

*Sarapin is a suspension of powdered sarracenia purpurea (pitcher plant). This study had a total of four patient groups, but outcomes were not reported by use of this product.

Harms

- Major complications are rare following injections into the cervical spine and included a life-threatening generalized anaphylactic reaction, grand-mal seizure, dural puncture, subarachnoid puncture, and local hematoma. There were no cases of death or paralysis in the included studies, although there have been case reports of each in the published literature.
- Multiple publications have addressed the outbreak in fungal infections and deaths that resulted from the use of contaminated steroid preparations when delivering epidural steroid injections. As of December 27, 2012, there had been 620 cases of fungal meningitis, stroke presumed due to fungal meningitis, other central nervous system-related infection or septic arthritis attributed to the contamination, and 39 deaths.
- Other major complications were reported in case series of a mixture of lumbar and cervical spinal injection patients and included chest pain, tachycardia/hypertension, significant transient hypertensive episode, hematoma, dural puncture, and a severe vasovagal reaction.
- Minor complications are more common following lumbar or cervical spinal injections but are generally transient in nature, and include pain at the injection site, increased radicular pain/numbness/weakness, nerve root irritation, superficial infections, sympathetic blockade, facial flushing or rash, vasovagal reactions/fainting, headache, gastric complaints, dizziness, pruritus, irregular periods, and insomnia.
- With proper protective measures, total radiation exposure to the physician was within normal limits following a mean of 923 procedures (range, 100 to 1819) with an average length of radiation exposure of 9.8 seconds/procedure (range, 4.9 to 15.2) in all five case series identified.

Table 2. Harms

Spinal Injections	Strength of Evidence	Conclusions
Major complications	High	Major complications are rare following injections into the cervical spine. There were no cases of death or paralysis in the included studies, although there have been case reports of each in the published literature. In five RCTs, there were reports of subarachnoid puncture in 3/710 injections or patients and no reports of dural puncture or death. In four case series there were reports of life-threatening generalized anaphylactic reaction (1 case), grand-mal seizure (1 case), dural puncture (2 cases), and local hematoma (1 case) in

Spinal Injections	Strength of Evidence	Conclusions
		7240 injections or patients. In three case reports of a mix of lumbar and cervical spinal injection patients, there was one case of each of the following major complications in 6935 injections: chest pain, tachycardia/hypertension, significant transient hypertensive episode, hematoma, dural puncture, and a severe vasovagal reaction.
Minor complications	High	Minor complications are more common but are generally transient in nature. They were rare as reported in 5 RCTs. One non-randomized study reported an overall minor complication rate of 1.64%. Complications included: pain at the injection site, increased usual pain, transient pain/weakness, nerve root irritation, transient global amnesia, superficial infections, sympathetic blockade, vasovagal reactions, facial flushing, headache, nausea.
Vascular puncture	Low	The mean incidence of intravascular puncture following fluoroscopically guided cervical spinal injections was 15.6% (range, 4.0–19.4%) as reported in two studies. These studies evaluated the incidence but not the consequences of intravascular injection. The TruCath Spinal Injection System™ may decrease the incidence of vascular puncture.
Radiation exposure to the physician	Low	With proper protective measures, total radiation exposure was within normal limits following a mean of 923 procedures (range, 100 – 1819) with an average length of radiation exposure of 9.8 seconds/procedure (range, 4.9 – 15.2) in all five case series identified.

For RF neurotomy, the only adverse event reported more frequently for RF neurotomy compared to sham was numbness in the area of the treated nerves. Other adverse events reported were psoriatic rash and pain associated with the procedure, neither of which differed significantly from the sham group. Overall strength of the evidence is low.

Additional Sources

Because of the weak evidence base for ESI, the subcommittee chose to consider additional evidence that was provided by the HERC-appointed expert for this topic. The expert provided a systematic review and a retrospective cohort study, both described below.

Cohen 2013

This review included an extensive search of the literature, and all study types were included. However, details of the search, including number and type of studies, were not provided. Authors state that “recommendations were based on a conglomeration of factors including weighted evidence ..., consensus guidelines when relevant, and perceived bias” (p. 175). No quality assessment of included studies was completed, and there was no explicit link between the evidence and recommendations. The review provides a summary of the literature and recommendations for use of epidural steroids in the spine.

Pertaining to injection approach:

There are two different approaches for delivering epidural steroids in the cervical region, interlaminar (IL) and transforaminal (TF). An interlaminar injection involves passage of the needle through the ligamentum flavum in the posterior midline of the spine. Advantages of this approach are purported to be a higher likelihood that the injected medication will reach adjacent spinal levels and the ability to treat bilateral pain. Disadvantages include the potential for dural puncture and the dorsal deposition of the medication, which is more distant from the site of pathology. The transforaminal approach involves placement of the needle within the neuroforamen. Theoretical advantages are that it is the most target specific, carries a lower risk of dural puncture and is associated with a greater incidence of ventral spread. However, they are associated with a higher risk of catastrophic neurological complications.

Pertaining to the efficacy of ESI in the cervical region:

For interlaminar injections, the authors cite three systematic reviews, one of which was supportive of the efficacy of ESI for a variety of conditions, while the other two reported that ESIs are probably effective in the short-term. One of these stated that definitive evidence was lacking, and that the evidence was stronger for herniated disc and non-osseous central stenosis. The other reported that evidence was lacking for long-term efficacy.

For transforaminal injections, the authors report a single RCT, but multiple nonrandomized studies reporting both short and long term efficacy. They also note multiple reports of potentially catastrophic complications. The authors recommend the following:

“Overall, the literature suggests that although the TF approach may be more efficacious than the IL or caudal approaches, the difference in effect size is small. In the cervical, thoracic, and midlumbar to high lumbar regions, the increased risk

for catastrophic neurological complications should preclude the use of TF ESI as a first-line treatment.” (p. 181)

Lee 2012

Lee (2012) is a retrospective cohort study in which 98 patients who received cervical ESI for cervical radiculopathy without major neurologic deficit were followed for two years, then were analyzed based on whether they received spine surgery or not. Patients must have failed 4 weeks of conservative therapy and not have myelopathy or definite motor weakness. The injection was repeated up to 3 times at 1 to 2 week intervals, if pain relief was insufficient (less than 80% improved). Authors report that patients received an average of 1.8 ESI, and that at the 2-year follow up point, 80.6% of patients had not had surgery.

Analysis of clinical factors associated with proceeding to surgery found that having had a previous episode of cervical radiculopathy, pre-ESI pain score and post-ESI pain score were the only correlated factors. Fifteen percent of patients in the non-surgery group had a previous episode of cervical radiculopathy, compared to 42% in the surgery group ($p=0.02$). Prior to ESI, patients in the non-surgery group had a mean visual analogue scale (VAS) score for arm pain of 6.1, compared to 8.2 in the surgical group. Post ESI, the non-surgical group had a mean VAS pain score of 2.8 compared to 6.9 for the surgical group ($p<0.001$ for both). Factors for which there was no statistically significant difference between groups included gender, age, duration of symptoms, type of symptoms, compensation or number of ESI.

Evidence Summary

There is no evidence of benefit of epidural steroid injections compared with placebo injections for neck pain in patients either with or without disc herniation and radiculitis, post-surgery syndrome or cervical spinal stenosis. On the other hand, epidural injections appear to be superior to intramuscular injections in patients with disc compression and radiculitis. Conclusions regarding the efficacy of ESI from other sources are mixed, but suggest if they are effective, it is likely only short-term. One study suggests that ESI may result in a decreased risk of cervical surgery. There is no evidence of benefit of cervical intraarticular facet joint steroid injections or medial branch blocks compared with placebo injections for confirmed facet joint pain. There is no apparent difference in efficacy based on differing approaches for administering cervical epidural steroids. There is limited evidence of benefit of RF neurotomy in patients with confirmed facet joint pain.

Major complications are rare following injections into the cervical spine but can include a life-threatening generalized anaphylactic reaction, grand-mal seizure, dural and subarachnoid puncture, paralysis and death. Expert opinion in one review notes multiple

reports of potentially catastrophic complications using a transforaminal injection approach and recommends against their use, despite similar or better efficacy compared to intralaminar approaches. An outbreak in fungal infections and deaths that resulted from the use of contaminated steroid preparations when delivering epidural steroid injections occurred in 2012. Minor complications are more common but are generally transient in nature. No major complications of RF neurotomy have been reported.

GRADE-INFORMED FRAMEWORK

The HERC develops recommendations by using the concepts of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system. GRADE is a transparent and structured process for developing and presenting evidence and for carrying out the steps involved in developing recommendations. There are four elements that determine the strength of a recommendation, as listed in the table below. The HERC reviews the evidence and makes an assessment of each element, which in turn is used to develop the recommendations presented in the coverage guidance box. Balance between desirable and undesirable effects, and quality of evidence, are derived from the evidence presented in this document, while estimated relative costs, values and preferences are assessments of the HERC members.

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variability in values and preferences	Coverage Recommendation	Rationale
Epidural steroid injections	Balance may result in some net benefit	Effectiveness evidence: Very low Harms evidence: High	Moderate	Moderate to high, given some may prefer interventional vs. pharmacologic vs. lifestyle or more conservative interventions	Recommended for coverage, only when all the following criteria are met: documented neuroforaminal stenosis (without infection or neoplasia); radicular pain in a corresponding dermatomal distribution; pain is intractable and conservative therapy has failed, fluoroscopic guidance or CT guidance is utilized; interlaminar approach is utilized; no more than two injections without clinically meaningful improvement in pain and function; and maximum of three injections in six months. (<i>weak recommendation</i>)	Though quality of evidence from trusted sources is very low, there is evidence from additional sources, namely a retrospective single cohort study, of some benefit when the recommended criteria are met. Additionally, other payer policies include a similar recommendation, namely Medicare and Washington State's payer policies.
Facet joint injections	No benefit	Effectiveness evidence: Very low Harms evidence:	Moderate	Moderate to high, given some may prefer interventional	Not recommended for coverage (<i>strong recommendation</i>)	No net benefit; some harms

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variability in values and preferences	Coverage Recommendation	Rationale
		High		vs. pharmacologic vs. lifestyle or more conservative interventions		
Therapeutic medial branch blocks	No benefit	Effectiveness evidence: Very low Harms evidence: High	Moderate	Moderate to high, given some may prefer interventional vs. pharmacologic vs. lifestyle or more conservative interventions	Not recommended for coverage (<i>strong recommendation</i>)	No net benefit; not common practice; some harms
RF neurotomy	Some evidence of benefit, minimal harms	Effectiveness evidence: Very low Harms evidence: High	Moderate	Moderate to high, given some may prefer interventional vs. pharmacologic vs. lifestyle or more conservative interventions	Recommended for coverage, only when all of the following criteria are met: at least 3 months of moderate to severe pain with functional impairment; pain is predominantly axial and not associated with radiculopathy; conservative therapy has failed; and at least 80% improvement from initial pain level following fluoroscopically-guided, low- volume local anesthetic blocks of the medial branch nerves, performed on two separate occasions. (<i>weak recommendation</i>)	Though quality of evidence from trusted sources is very low, the evidence shows there is some benefit when the recommended criteria are met. Coverage with these criteria was recommended by appointed expert.

*The Quality of Evidence rating was assigned by the primary evidence source, not the HERC Subcommittee

Note: GRADE framework elements are described in Appendix A

POLICY LANDSCAPE

Quality Measures

No pertinent quality measures were identified when searching the [National Quality Measures Clearinghouse](#).

Professional Society Guidelines

Guidelines reviewed in the WA HTA Spinal Injections report (Hashimoto, 2011) are included in Appendix D.

Payer Coverage Policies

Coverage policies for selected payers are included here.

Medicare

Two local coverage determinations were identified in the [Medicare Coverage Database](#).

- [L30481 Epidural and Transforaminal Epidural Injections](#)
- [L33842 Facet Joint Injections, Medial Branch Blocks, and Facet Joint Radiofrequency Neurotomy](#)

Medicare coverage guidance is detailed in Appendix E.

Washington HTA Limitations of Coverage

Therapeutic Epidural Injections in the lumbar or cervical-thoracic spine for chronic pain are a covered benefit when all of the following conditions are met:

- For treatment of radicular pain
- With fluoroscopic guidance or CT guidance
- After failure of conservative therapy
- No more than two without clinically meaningful improvement in pain and function, and
- Maximum of 3 in 6 months

COMMITTEE DELIBERATIONS – VALUE-BASED BENEFITS SUBCOMMITTEE

Coverage guidance is prepared by the Health Evidence Review Commission (HERC), HERC staff, and subcommittee members. The evidence summary is prepared by the Center for Evidence-based Policy at Oregon Health & Science University (the Center). This document is intended to guide public and private purchasers in Oregon in making informed decisions about health care services.

The Center is not engaged in rendering any clinical, legal, business or other professional advice. The statements in this document do not represent official policy positions of the Center. Researchers involved in preparing this document have no affiliations or financial involvement that conflict with material presented in this document.

DRAFT

Appendix A. GRADE Element Descriptions

Element	Description
Balance between desirable and undesirable effects	The larger the difference between the desirable and undesirable effects, the higher the likelihood that a strong recommendation is warranted. The narrower the gradient, the higher the likelihood that a weak recommendation is warranted
Quality of evidence	The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted
Resource allocation	The higher the costs of an intervention—that is, the greater the resources consumed—the lower the likelihood that a strong recommendation is warranted
Values and preferences	The more values and preferences vary, or the greater the uncertainty in values and preferences, the higher the likelihood that a weak recommendation is warranted

Strong recommendation

In Favor: The subcommittee is confident that the desirable effects of adherence to a recommendation outweigh the undesirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences.

Against: The subcommittee is confident that the undesirable effects of adherence to a recommendation outweigh the desirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences.

Weak recommendation

In Favor: The subcommittee concludes that the desirable effects of adherence to a recommendation probably outweigh the undesirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences, but is not confident.

Against: The subcommittee concludes that the undesirable effects of adherence to a recommendation probably outweigh the desirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences, but is not confident.

Quality or strength of evidence rating across studies for the treatment/outcome¹

High: The subcommittee is very confident that the true effect lies close to that of the estimate of the effect. Typical sets of studies are RCTs with few or no limitations and the estimate of effect is likely stable.

Moderate: The subcommittee is moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Typical sets of studies are RCTs with some limitations or well-performed nonrandomized studies with additional strengths that guard against potential bias and have large estimates of effects.

Low: The subcommittee's confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. Typical sets of studies are RCTs with serious limitations or nonrandomized studies without special strengths.

Very low: The subcommittee has very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect. Typical sets of studies are nonrandomized studies with serious limitations or inconsistent results across studies.

¹ Includes risk of bias, precision, directness, consistency and publication bias

Appendix B. Applicable Codes

CODES	DESCRIPTION
ICD-9 Diagnosis Codes	
720.9	Unspecified inflammatory spondylopathy
721	Spondylosis , various
722	Degeneration/displacement intervertebral disc, various
723	Spinal stenosis, brachial neuritis or radiculitis, various
724	Spinal stenosis, various
738.4	Acquired spondylolisthesis
ICD-10 Diagnosis Codes	
M43.00-3	Spondylolysis
M43.10-3	Spondylolisthesis
M47.10-3	Other spondylosis with myelopathy
M47.20-3	Other spondylosis with radiculopathy
M47.811-3, M47.819	Spondylosis without myelopathy or radiculopathy
M47.891-3	Other spondylosis
M48.00-3	Spinal stenosis
M50	Cervical disc disorders
M54.10-3	Radiculopathy
M54.2	Cervicalgia
ICD-9 Volume 3 (Procedure Codes)	
03.91	Injection of anesthetic into spinal canal for analgesia
03.92	Injection of another agent into spinal canal
CPT Codes	
62310	Inject spine cervical/thoracic
64479	Inj foramen epidural c/t
64480	Inj foramen epidural add-on
64490- 64495	Injection(s), diagnostic or therapeutic agent, paravertebral facet (zygapophyseal) joint (or nerves innervating that joint) with image guidance (flouro or CT), cervical, thoracic, lumbar or sacral; single, second or third level
64633- 64636	Destruction by neurolytic agent, paravertebral facet (zygapophyseal) joint (or nerves innervating that joint) with image guidance (flouro or CT), cervical, thoracic, lumbar or sacral; single, second or third level
64479-80	Injection(s), anesthetic agent and/or steroid, transforaminal epidural, with imaging guidance; cervical or thoracic, single level and each additonal
64490-2	Injection(s), diagnostic or therapeutic agent, paravertebral facet joint (or nerves innervating that joint), with imaging guidance; cervical or thoracic, single level and each additional
0213T– 0218T	Injection(s), diagnostic or therapeutic agent, paravertebral facet (zygapophyseal) joint (or nerves innervating that joint) with ultrasound guidance, cervical, thoracic, lumbar or sacral; single, second or third level
77003	Fluoroguide for spine injection
HCPCS Level II Codes	
None	

Note: Inclusion on this list does not guarantee coverage

Appendix C. HERC Guidance Development Framework

HERC Guidance Development Framework Principles

This framework was developed to assist with the decision making process for the Oregon policy-making body, the HERC and its subcommittees. It is a general guide, and must be used in the context of clinical judgment. It is not possible to include all possible scenarios and factors that may influence a policy decision in a graphic format. While this framework provides a general structure, factors that may influence decisions that are not captured on the framework include but are not limited to the following:

- Estimate of the level of risk associated with the treatment, or any alternatives;
- Which alternatives the treatment should most appropriately be compared to;
- Whether there is a discrete and clear diagnosis;
- The definition of clinical significance for a particular treatment, and the expected margin of benefit compared to alternatives;
- The relative balance of benefit compared to harm;
- The degree of benefit compared to cost; e.g., if the benefit is small and the cost is large, the committee may make a decision different than the algorithm suggests;
- Specific indications and contraindications that may determine appropriateness;
- Expected values and preferences of patients.

Facet Joint Injections, Therapeutic Medial Branch Blocks



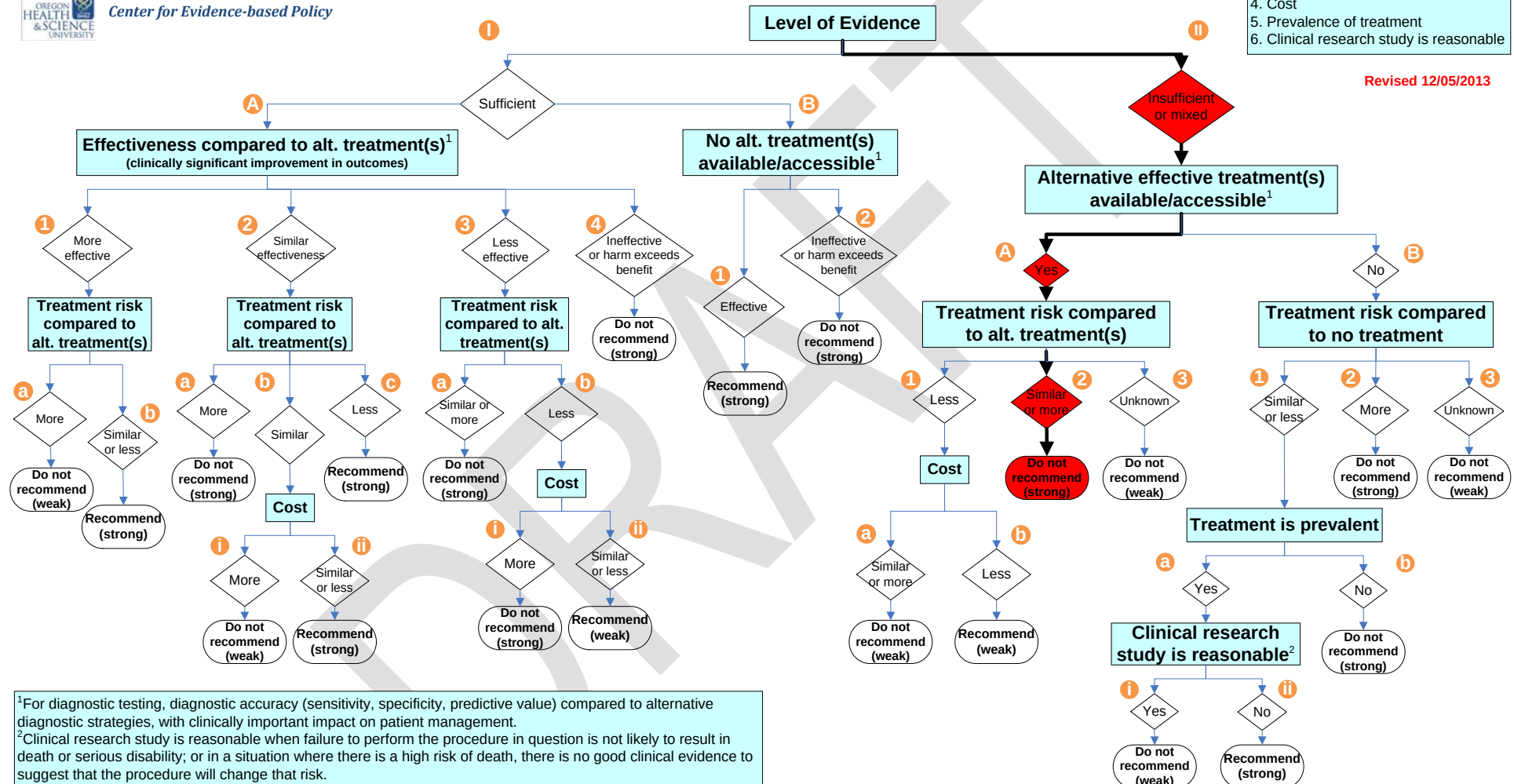
Center for Evidence-based Policy

HERC Guidance Development Framework

Refer to *HERC Guidance Development Framework Principles* for additional considerations

- Decision Point Priorities**
1. Level of evidence
 2. Effectiveness & alternative treatments
 3. Harms and risk
 4. Cost
 5. Prevalence of treatment
 6. Clinical research study is reasonable

Revised 12/05/2013



Epidural Steroid Injections, RF Neurotomy



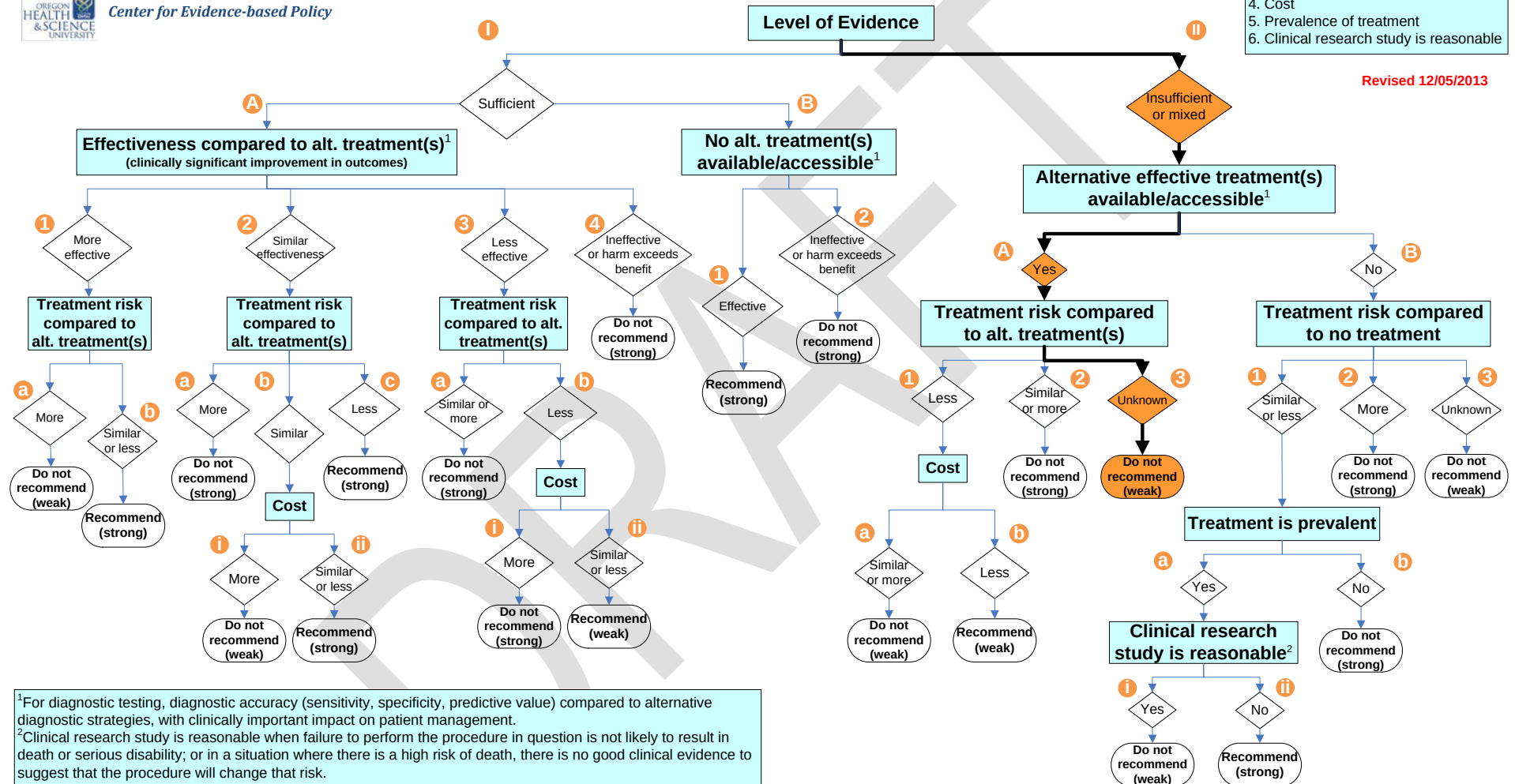
Center for Evidence-based Policy

HERC Guidance Development Framework

Refer to HERC Guidance Development Framework Principles for additional considerations

- Decision Point Priorities**
1. Level of evidence
 2. Effectiveness & alternative treatments
 3. Harms and risk
 4. Cost
 5. Prevalence of treatment
 6. Clinical research study is reasonable

Revised 12/05/2013



¹For diagnostic testing, diagnostic accuracy (sensitivity, specificity, predictive value) compared to alternative diagnostic strategies, with clinically important impact on patient management.

²Clinical research study is reasonable when failure to perform the procedure in question is not likely to result in death or serious disability; or in a situation where there is a high risk of death, there is no good clinical evidence to suggest that the procedure will change that risk.

Appendix D. Guidelines Included in the WA HTA Spinal Injections Report

Hashimoto, R., Raich, A., Ecker, E., Henrikson, N., Wallace, L., Dettori, J., & Chou, R. (2011). *Spinal Injections*. Olympia, WA: Washington Health Technology Assessment Program. Retrieved from http://www.hta.hca.wa.gov/documents/spinal_injection_draft_report.pdf

Fourteen guidelines were reported in this report. Of those, five pertain to cervical pain. One of those addressed chronic pain in general, was from 2008 and could not be retrieved, therefore was not assessed further. At the request of HERC staff, the quality of the other four guidelines was rated using an instrument adapted from the Appraisal of Guidelines Research and Evaluation (AGREE) Collaboration. A summary of the pertinent recommendations and the quality rating of each guideline is presented in the table below:

Guideline	Recommendations	Quality
American Society of Interventional Pain Physicians. (2009) Comprehensive evidence-based guidelines for interventional techniques in the management of chronic spinal pain (NGC:007428). [Updated in 2013]. http://asipp.org/Guidelines.htm	Cervical discography is indicated only when a treatment is available to test the diagnostic hypothesis of discogenic pain of the cervical spine in individuals who have been properly selected and screened to eliminate other sources of cervical pain. There is good evidence for disc herniation and fair evidence for axial or discogenic pain, central spinal stenosis, and post cervical surgery syndrome. Cervical interlaminar epidural injections are indicated for these conditions with appropriate indications. Diagnostic cervical facet joint nerve blocks are recommended in patients with somatic or non-radicular neck pain or headache and upper extremity pain, with duration of pain of at least 3 months, without preponderance of evidence of discogenic pain, disc herniation, or evidence of radiculitis. There is fair evidence for conventional radiofrequency neurotomy and therapeutic facet joint nerve blocks with limited evidence for intraarticular injections. Consequently, the recommendation is that therapeutic facet joint nerve blocks or conventional radiofrequency neurotomy may be provided based on the response from controlled diagnostic blocks.	Poor
Institute for Clinical Systems Improvement. (2009). <i>Assessment and management of chronic pain</i> (NGC:007602). [Updated in 2013]. https://www.icsi.org/guidelines_m	Examples of commonly used therapeutic procedures are as follows. Facet joint injection Facet joints are an important source of spinal pain in the cervical and lumbar regions. These joints can be reliably anesthetized by way of fluoroscopically guided joint injections. Generally, a depot corticosteroid is administered concomitantly, which may provide short-term benefit for a subset of patients. However, clinical trials have failed to demonstrate any sustained therapeutic benefits following facet joint corticosteroid injections (<i>Nelemans, 2005 [Systematic Review]</i>).	Poor

Guideline	Recommendations	Quality
ore/catalog_guidelines_and_more/catalog_guidelines/catalog_neurological_guidelines/pain/	Percutaneous radiofrequency neurotomy Percutaneous radiofrequency (RF) neurotomy (sometimes erroneously referred to as facet rhizotomy) is a treatment for neck or back pain generated by facet joints. Properly selected candidates for this procedure should experience complete or nearly complete relief of their pain following fluoroscopically guided, low-volume local anesthetic blocks of the medial branch nerves that innervate the pain-generating joint(s). To minimize false-positive results, an equivalent degree of relief of appropriate pharmacologic duration should be carefully documented on two separate occasions, using two different types of local anesthetic. The radiofrequency procedure is performed by placing an insulated needle electrode with an exposed tip adjacent to and in parallel with the medial branch nerves that supply the target joint(s). Radiofrequency current applied to the electrode then heats the adjacent tissues and coagulates the nerve supply to the joint. For the procedure to be effective, multiple lesions must be performed at each nerve location, using electrodes of sufficient diameter. The nerves do regenerate over time, so pain relief is not permanent, but the procedure can be repeated. Radiofrequency neurotomy can provide pain relief for carefully selected patients, but this procedure should be performed only by an experienced pain medicine physician in the context of a longitudinal and comprehensive care plan. Proper patient selection and appropriate technique in positioning the radiofrequency electrodes are absolutely essential to the success of the procedure (<i>Bogduk, 2008 [Low Quality Evidence]; Nath, 2008 [Moderate Quality Evidence]; Hooten, 2005 [Guideline]</i>). Controversy in the literature regarding the efficacy of lumbar radiofrequency neurotomy has arisen from fundamentally flawed clinical trials that have used inappropriate patient selection criteria, and improper procedural technique. Epidural corticosteroid injections Epidural corticosteroid injections are one of the most commonly performed interventions for treatment of spinal pain with a radicular component. All epidural injections should be performed by an experienced physician, under fluoroscopic guidance, using contrast injection to detect vascular uptake and to demonstrate the injectate spread pattern. There are three approaches to the epidural space, including a transforaminal, intralaminar and a caudal technique. Limited evidence was found to support the efficacy of this procedure (<i>Riew, 2000 [High Quality Evidence]; Carette, 1997 [High Quality Evidence]; Dilke, 1973 [High Quality Evidence]</i>). Transforaminal epidural injection Transforaminal epidural injections can be used to determine the spinal level that is the source of radicular pain. The risks of cervical transforaminal epidural steroid injections have been well documented in case reports (<i>Beckman, 2006 [Low Quality Evidence]; Tiso, 2004</i>	

Guideline	Recommendations	Quality
	<i>[Low Quality Evidence]; Furman, 2003 [Low Quality Evidence]</i>). Specifically, cervical transforaminal epidural steroid injections have been associated with spinal cord and brain injuries resulting in permanent neurological deficits and/or death. These adverse events have been caused by uptake of particulate corticosteroids into radicular or vertebral arteries, producing embolization, severe vasospasm, and either brain or spinal cord infarction. For cervical procedures, it is recommended that only non-particulate corticosteroids be utilized. These procedures should be performed only by an experienced pain medicine physician with access to and knowledge of the use of appropriate imaging equipment and patient monitoring facilities, and should be performed only in the context of a longitudinal care plan, as directed and coordinated by a pain medicine physician (<i>Bogduk, 2008 [Low Quality Evidence]; Tiso, 2004 [Low Quality Evidence]</i>).	
Work Loss Data Institute. (2008). <i>Neck and upper back (acute & chronic)</i> (NGC:006563). [Updated 2011]. Unable to retrieve full guideline (subscription only). http://www.guideline.gov/content.aspx?id=33185	This guideline was not available without full subscription to the WLDI. Review of the summary on the NGC recommends consideration of ESI on the 4th visit, 3-4 weeks after onset of pain, only in those with neurologic findings, only for those “severe cases hoping to avoid surgery”.	Unable to assess quality without access to full guideline
Work Loss Data Institute. (2008). <i>Pain (chronic)</i> (NGC:006564). [Updated 2011]. Unable to retrieve full guideline (subscription only). http://www.guideline.gov/content.aspx?id=33188	This guideline was not available without full subscription to the WLDI. Review of the summary on the NGC did not specifically address neck pain, but states that epidural steroid injections and facet blocks are recommended. Additional information about indications is not provided in the NGC summary.	Unable to assess quality without access to full guideline

Appendix E. CMS Local Coverage Decisions

L30481 Epidural and Transforaminal Epidural Injections

Original Effective Date: For services performed on or after 04/15/2010

Revision Effective Date: For services performed on or after 02/01/2014

Jurisdiction: Includes Oregon

Coverage Indications, Limitations, and/or Medical Necessity

Epidural injections are used for acute and chronic pain, in addition to cancer pain management. Epidural injections are utilized both for diagnostic and therapeutic purposes.

A multi-disciplinary or collaborative comprehensive evaluation (e.g. orthopedics, neurologist, neurosurgeon, physiatrist, anesthesiologist, pain medicine specialist, and/or attending physician), is recommended prior to initiating a trial of these injections for relief of chronic recurrent pain.

Epidural steroid injections, both interlaminar/translaminar and transforaminal should be used only in the presence of radiculopathy.

Indications for Diagnostic and Therapeutic Epidural Injections

Diagnostic interlaminar/translaminar or caudal epidural steroid injections are seldom used. Although the medication injected can sometimes be confined to a limited area, bilateral effects and spread to adjacent levels often occur.

Therapeutic interlaminar/translaminar or caudal epidural injections and infusions of opioid, local anesthetic, or other medications may be used for the treatment of acute and chronic pain or cancer pain.

Epidural injections (interlaminar/translaminar or caudal) may be used for the following.

- Acute obstetric, post traumatic and postoperative pain
- Advanced cancer pain, primary or metastatic
- Acute/sub acute and chronic pain syndrome including cervical, thoracic and lumbar pain with radiculopathy and intervertebral disc disease (with neuritis or radiculitis) with or without myelopathy that has failed to respond to adequate conservative management.
- Nerve root injuries and neuropathic pain and post traumatic including post laminectomy syndrome (failed back syndrome).
- Spinal cord myelopathy
- Complex regional pain syndrome
- Epidural scarring from prior infection, hemorrhage and/or surgery

- Multiple rib fractures
- Vertebral compression fractures
- Post herpetic neuralgia and herpes zoster
- Phantom limb pain

Indications for Diagnostic and Therapeutic Transforaminal Epidural Injections

Transforaminal epidural injection is a selective injection of the cervical, thoracic, lumbar or sacral nerve roots with proximal spread of contrast or local anesthetic through the neural foramen to the epidural space. With the aid of fluoroscopic or computed tomography (CT) imaging, the needle tip is placed within or adjacent to the lateral margin of the neural foramen and contrast material is injected to obtain a neurogram and visualize spread of the injected solution.

A small volume of local anesthetic is injected (less than or equal to 1.0 ml) in order to perform a diagnostic reproducible blockade of a specific nerve root. The diagnostic usefulness is lost if more than 1.0 ml of local anesthetic is injected (the block becomes unreliable since the spread of anesthetic to adjacent levels and structures likely occurs).

Diagnostic transforaminal epidural injections are appropriate for the following purposes.

- To differentiate the level of radicular nerve root pain.
- To differentiate radicular from non radicular pain
- To evaluate a discrepancy between imaging studies and clinical findings
- To identify the source of pain in the presence of multi-level nerve root compression
- To identify the level of pathology at a previous operative site

It might be necessary to perform injections at two different nerve root levels on the same date of service. When multiple levels of nerve root compression or stenosis is suspected to be responsible for the patient's symptoms, presence of the compression or stenosis on imaging studies should be documented in the medical record.

Therapeutic transforaminal epidural injections are appropriate for the following purposes:

Corticosteroid can be added as a therapeutic measure. Injections for therapeutic reasons can be of greater volume. The transforaminal injection can be performed for diagnostic, therapeutic or both purposes.

- Radicular pain resistant to more conservative measures or when surgery is contraindicated.
- Post-decompressive radiculitis or post surgical scarring

- Monoradicular pain, confirmed by diagnostic block in which a surgically correctible lesion cannot be identified
- Treatment of acute herpes zoster or post herpetic neuralgia

General Indications and Limitations

Epidural (interlaminar/translaminar or caudal) and transforaminal epidural corticosteroid injections should not exceed a series of three, per spinal region, within a six-month period when used as treatment for a pain disorder other than treatment for cancer pain. These may be performed at intervals of one week or greater. With each subsequent injection the medical record should clearly document the interval effects from the prior injection(s). Appropriate reasons for a repeat injection are: (a) significant improvement in the patient's symptoms from the prior injection, even if relapsed, or (b) carefully documented technical reasons that it is appropriate to repeat the procedure even if no prior improvement and (c) patients with persistent pain in whom the imaging findings suggest that the pathology should respond to corticosteroid injection. In the absence of a compelling technical reason, it is not appropriate to repeat a procedure a third time if there has been no improvement from the two preceding.

If corticosteroids are used, consideration should be given to the potential complications of repetitive corticosteroid administration.

Many of these procedures, such as those in the peri-operative period, may not require fluoroscopy.

For treatment of chronic pain, the standard of care is that these procedures be performed under fluoroscopic or CT guided imaging. Therefore injections for chronic pain performed without imaging guidance will be considered not medically necessary.

Fluoroscopic guidance **must** be utilized in the performance of single nerve root/transforaminal injections to ensure the precise placement of the needle and medications injected.

Anti-spasmodic drugs administered intrathecally (e.g., baclofen) to treat chronic intractable spasticity are addressed in the Infusion Pump NCD Pub. 100-3 Sec. 280.14. The CPT description of procedure codes 62310, 62311, 62318 and 62319 include anesthetic, antispasmodic, opioid, steroid, other solution; therefore the spasticity conditions are included in this LCD.

L33842 Facet Joint Injections, Medial Branch Blocks, and Facet Joint Radiofrequency Neurotomy

Original Effective Date: For services performed on or after 03/05/2014

Revision Effective Date: For services performed on or after 03/05/2014

Jurisdiction: Oregon

Coverage Indications, Limitations, and/or Medical Necessity

Introduction:

This policy does not address sacral conditions or injections or neurotomies. Sacral injections, identified on the claim by the ICD-9 code 724.6, are not subject to the requirements of this LCD.

Facet joints are paired diarthrodial articulations of the superior and inferior articular processes of adjacent vertebrae. The medial branches (MB) of the dorsal rami of the segmental nerves innervate facet joints and the MB nerves from the two adjacent dorsal rami innervate each joint. [Exceptions to this rule are the C2-3 facet joint, which is innervated by the third occipital nerve; and the L5-S1 facet joint, which is innervated by the L4 MB and the L5 dorsal ramus.]

Facet joint injection techniques are used in the diagnosis and/or treatment of chronic neck and back pain. However, the evidence of clinical efficacy and utility has not been well-established in the medical literature, which is replete with non-comparable and inadequately designed studies. Further, there is a singular dearth of long-term outcomes reports. This is particularly problematic given the steroid dosages administered. These drugs alone may develop the relief experienced by patients but are associated with serious adverse health events and could as well be administered orally. Hence, ongoing coverage requires outcomes reporting as described in this LCD to allow future analysis of clinical efficacy.

Definitions

- A zygapophyseal (aka facet) joint “level” refers to the zygapophyseal joint or the two medial branch (MB) nerves that innervate that zygapophyseal joint.
- A “session” is defined as all injections/blocks/RF procedures performed on one day and includes medial branch blocks (MBB), intraarticular injections (IA), facet cyst ruptures, and RF ablations.
- A “region” is all injections performed in cervical/thoracic or all injections performed in lumbar (not sacral) spinal areas.
- “Diagnosis” of facet-mediated pain requires the establishment of pain relief following dual medial branch blocks (MBBs) performed at different sessions.

Neither physical exam nor imaging has adequate diagnostic power to confidently distinguish the facet joint as the pain source.

Indications

- Patient must have history of at least 3 months of moderate to severe pain with functional impairment and pain is inadequately responsive to conservative care such as NSAIDs, acetaminophen, physical therapy (as tolerated).
- Pain is predominantly axial and not associated with radiculopathy or neurogenic claudication.
- There is no non-facet pathology that could explain the source of the patient's pain, such as fracture, tumor, infection, or significant deformity.
- Clinical assessment implicates the facet joint as the putative source of pain.

General Procedure Requirements:

- Pre-procedural documentation must include a complete initial evaluation including history and an appropriately focused musculoskeletal and neurological physical examination. There should be a summary of pertinent diagnostic tests or procedures justifying the possible presence of facet joint pain.
- A procedure note must be legible and include sufficient detail to allow reconstruction of the procedure. Required elements of the note include a description of the techniques employed, nerves injected and sites(s) of injections, drugs and doses with volumes and concentrations as well as pre and post-procedural pain assessments. With RF neurotomy, electrode position, cannula size, lesion parameters, and electrical stimulation parameters and findings must be specified and documented.
- Facet joint interventions (diagnostic and/or therapeutic) must be performed under fluoroscopic or computed tomographic (CT) guidance. Facet joint interventions performed under ultrasound guidance will not be reimbursed.
- A hard (plain radiograph with conventional film or specialized paper) or digital copy image or images which adequately document the needle position and contrast medium flow (excluding RF ablations and those cases in which using contrast is contra-indicated, such as patients with documented contrast allergies), must be retained and submitted if requested.
- In order to maintain target specificity, total IA injection volume must not exceed 1.0 mL per cervical joint or 2 mL per lumbar joint, including contrast. Larger volumes may be used only when performing a purposeful facet cyst rupture in the lumbar spine.
- Total MBB anesthetic volume shall be limited to a maximum of 0.5 mL per MB nerve for diagnostic purposes and 2ml for therapeutic. For a third occipital nerve block, up to 1.0 mL is allowed for diagnostic and 2ml for therapeutic purposes.

- In total, no more than 100 mg of triamcinolone or methylprednisolone or 15 mg of betamethasone or dexamethasone or equivalents shall be injected during any single injection session.
- Both diagnostic and therapeutic facet joint injections may be acceptably performed without steroids.

Provider Qualifications

Provider Qualifications' requirements must be met. Patient safety and quality of care mandate that healthcare professionals who perform Facet Joint Injections, Medial Branch Blocks, and Facet Joint Radiofrequency Neurotomy are appropriately experienced and/or trained to provide and manage the services. The CMS Manual System, Pub. 100-8, Program Integrity Manual, Chapter 13, Section 5.1 (<http://www.cms.hhs.gov/manuals/downloads/pim83c13.pdf>) underscores this point and states that "reasonable and necessary" services must be "ordered and/or furnished by qualified personnel." Services will be considered medically reasonable and necessary only if performed by appropriately experienced and/or formally trained providers.

The following training requirement applies only to those providers who have not provided these specific interventional pain management services on a regular basis (at least two times per month) during the ten years prior to the effective date of this LCD as may be established by claims billings. A basic requirement of payment is training and/or credentialing by a formal residency/fellowship program and/or other training program that accredited by a nationally-recognized body and whose core curriculum includes the performance and management of the procedures addressed in this policy. Recognized accrediting bodies include only those whose program accreditation gains the trainee eligibility to sit for a healthcare-related licensing exam or licensing itself, which in turn allows the licensee to perform these procedures. At a minimum, training must cover and develop an understanding of anatomy and drug pharmacodynamics and kinetics, the technical performance of the procedure(s) and utilization of the required associated imaging modalities, and the diagnosis and management of potential complications from the intervention.

The following credentialing requirement applies to all providers of the services addressed in this policy. If the practitioner works in a hospital facility at any time and/or is credentialed by a hospital for any procedure, the practitioner must be credentialed to perform the same procedure in the outpatient setting.

Diagnostic Facet Joint Injections

- Dual MBBs are necessary to diagnose facet pain due to the unacceptably high false positive rate of single MBB injections.

- A second confirmatory MBB is allowed if documentation indicates the first MBB produced $\geq 80\%$ relief of primary (index) pain and duration of relief is consistent with the agent employed.
- Intraarticular facet block will not be reimbursed as a diagnostic test unless medial branch blocks cannot be performed due to specific documented anatomic restrictions.

Therapeutic Injections

- Medial Branch Blocks may provide temporary or long-lasting or permanent relief of facet-mediated pain. Injections may be repeated if the first MBB results in significant pain relief ($> 50\%$) for at least 3 months. (See Limitations section for total number of injections that may be performed in one year.)
- Intraarticular injections may be covered for treatment of defined facet pain a) above or below a posterior spinal fusion when technical performance of MBBs is precluded, and/or b) when thermal RF neurotomy is precluded due to an implantable spinal cord stimulator or cardiac pacemaker or c) for rupture of symptomatic synovial cyst. Injections for these conditions and for axial pain with an arthritic joint may be repeated if the first intraarticular injection results in significant pain relief ($> 50\%$) for at least 3 months. (See Limitations section for total number of injections that may be performed in one year.)
- Recurrent pain at the site of previously diagnosed facet pain (dual MBBs) may be treated without additional diagnostic blocks if $> 50\%$ pain relief from the previous blocks lasted at least 3 months.

Thermal Medial Branch Radiofrequency Neurotomy (includes RF and microwave technologies):

- Only when dual MBBs provide $\geq 80\%$ relief of the primary or index pain and duration of relief is consistent with the agent employed may facet joint denervation with RF medial branch neurotomy be considered.
- Repeat denervation procedures involving the same joint will only be considered medically necessary if the patient experienced $\geq 50\%$ improvement of pain and improvement in patient specific ADLs documented for at least 6 months.

Limitations of Coverage:

- A maximum of five (5) facet joint injection sessions inclusive of medial branch blocks, intraarticular injections, facet cyst rupture and RF ablations may be performed per year in the cervical/thoracic spine and five (5) in the lumbar spine.
- For each covered spinal region (cervical/thoracic or lumbar), no more than two (2) thermal RF sessions will be reimbursed in any calendar year, involving no more than four (4) joints per session, e.g., two (2) bilateral levels or four (4) unilateral levels.

- Neither conscious sedation nor Monitored Anesthesia Care (MAC) is routinely necessary for intraarticular facet joint injections or medial branch blocks and are not routinely reimbursable. Individual consideration may be given for payment in rare unique circumstances if the medical necessity of sedation is unequivocal and clearly documented.
- Non-thermal RF modalities for facet joint denervation including chemical, low grade thermal energy (<80 degrees Celsius), as well as pulsed RF are not covered.
- Intraarticular and/or extraarticular facet joint prolotherapy is not covered.

COVERAGE GUIDANCE: ABLATION FOR ATRIAL FIBRILLATION

Question: How should the EGBS Coverage Guidance regarding ablation for atrial fibrillation be applied to the Prioritized List?

Question source: Evidence Based Guideline Subcommittee

Issue: EGBS approved a new Coverage Guidance at their September, 2014 meeting. This CG needs final approval by HERC at the November, 2014 meeting. The summary of the coverage guidance is shown below.

HERC COVERAGE GUIDANCE

AV node ablation is recommended for coverage only in persons with inadequate ventricular rate control resulting in symptoms, left ventricular systolic dysfunction or substantial risk of left ventricular systolic dysfunction. Coverage is recommended only when pharmacological therapy for rate control is ineffective or not tolerated (*weak recommendation*)

Transcatheter pulmonary vein isolation is recommended for coverage for those who remain symptomatic from atrial fibrillation despite rate control medications and antiarrhythmic medications (*strong recommendation*)

Surgical ablation (pulmonary vein isolation or Maze procedure) for atrial fibrillation is recommended for coverage at the time of other cardiac surgery for patients who remain symptomatic despite rate control medications (*weak recommendation*).

Evidence Summary

Ablation of the AV node or bundle of His in patients with AF results in lower heart rate at 12 months than pharmacologic treatment (moderate SOE), although there is no difference in mortality or exercise capacity (low SOE). Pulmonary vein isolation (PVI) results in a greater likelihood of maintaining sinus rhythm at 12 months than pharmacologic treatment (high SOE); most of the evidence for this finding is in patients with AF who have failed at least one AAD. This procedure (PVI) also results in lower risk of hospitalization over 12 months (moderate SOE) and improved QOL (moderate SOE), but the evidence is insufficient to assess the impact of PVI on mortality.

The surgical Maze procedure, when done at the time of other cardiac surgery, results in a higher likelihood of maintaining sinus rhythm than not performing the Maze (moderate SOE). Similarly, PVI done at the time of other cardiac surgery results in a higher likelihood of maintaining sinus rhythm than not performing PVI (high SOE), and no apparent difference in all-cause mortality or stroke (low SOE).

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variability in values and preferences	Coverage recommendation	Rationale
Ablation of AV node/bundle of His compared to rate control medications in patients for whom initial pharmacotherapy was ineffective	Lower heart rate, no difference in mortality/exercise capacity	Moderate/Low based on 1 to 3 poor to good quality studies, depending on the outcome	High	High	AV node ablation is recommended for coverage only in symptomatic persons when pharmacological therapy for rate control is ineffective or not tolerated. (<i>weak recommendation</i>)	Studies show mixed clinical significance of a lower heart rate. In those with persistently uncontrolled heart rate despite AADs, AV node ablation is a reasonable alternative to prevent the negative consequences of an uncontrolled rate such as MI, exacerbation of CHF or cardiomyopathy.
Transcatheter PVI vs. AAD	Better maintenance of SR, fewer hospitalizations, better QOL, possibly lower mortality	High/Moderate, based on 1 to 8 fair to good quality studies, depending on the outcome	High	Moderate	Transcatheter PVI is recommended for coverage when a rhythm control strategy is desired (<i>strong recommendation</i>)	Transcatheter PVI produces superior clinical outcomes to antiarrhythmic drugs alone when a rhythm control strategy is pursued
Maze procedure	Better maintenance of SR; possible (nonsignificant) increase in mortality	Moderate based on 1 good and six fair quality studies	Moderate (concurrent with other cardiac surgery)	Moderate	The Maze procedure is recommended for coverage at the time of other cardiac surgery if the benefits of maintenance of sinus rhythm are thought to	Maze may help maintain sinus rhythm but concerning nonsignificant increased risk of mortality

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variability in values and preferences	Coverage recommendation	Rationale
					outweigh the potential risk of increased mortality (<i>weak recommendation</i>)	
PVI done with other cardiac surgery	Better maintenance of SR	High based on 5 good and 3 fair quality studies	Moderate (concurrent with other cardiac surgery)	Low	PVI is recommended for coverage (<i>weak recommendation</i>)	PVI may help maintain sinus rhythm without significant additional risks

Current Prioritized List status

ICD-9 427.31 Atrial fibrillation is on line 350 CARDIAC ARRHYTHMIAS

CPT code	Code description	Current lines	Suggested lines
33250	Operative ablation of supraventricular arrhythmogenic focus or pathway (eg, Wolff-Parkinson-White, atrioventricular node re-entry), tract(s) and/or focus (foci); without cardiopulmonary bypass	286 LIFE-THREATENING CARDIAC ARRHYTHMIAS 350 CARDIAC ARRHYTHMIAS	286, 350
33251	...with cardiopulmonary bypass	286, 350	286, 350
33254	Operative tissue ablation and reconstruction of atria, limited (eg, modified maze procedure)	286, 350	350
33255	Operative tissue ablation and reconstruction of atria, extensive (eg, maze procedure); without cardiopulmonary bypass	286, 350	350
33256	...with cardiopulmonary bypass	286, 350	350
33257	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), limited (eg, modified maze procedure)	286, 350	350
33258	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), extensive (eg, maze procedure), without cardiopulmonary bypass	286, 350	350
33259	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), extensive (eg, maze procedure), with cardiopulmonary bypass	286, 350	350
33261	Operative ablation of ventricular arrhythmogenic focus with cardiopulmonary bypass	73 ACUTE AND SUBACUTE ISCHEMIC HEART DISEASE, MYOCARDIAL INFARCTION 193 CHRONIC ISCHEMIC HEART DISEASE 286,350	286,350
33265	Endoscopy, surgical; operative tissue ablation and reconstruction of atria, limited (eg, modified maze procedure), without cardiopulmonary bypass	286, 350	350
33266	...operative tissue ablation and reconstruction of atria, extensive (eg, modified maze procedure), without cardiopulmonary bypass	286, 350	350

93613	Intracardiac electrophysiologic 3-dimensional mapping	286, 350	286, 350
93650	Intracardiac catheter ablation of atrioventricular node function, atrioventricular conduction for creation of complete heart block, with or without temporary pacemaker placement	286, 350	286, 350
93653	Comprehensive electrophysiologic evaluation including insertion and repositioning of multiple electrode catheters with induction or attempted induction of an arrhythmia with right atrial pacing and recording, right ventricular pacing and recording, His recording with intracardiac catheter ablation of arrhythmogenic focus; with treatment of supraventricular tachycardia by ablation of fast or slow atrioventricular pathway, accessory atrioventricular connection, cavo-tricuspid isthmus or other single atrial focus or source of atrial re-entry	286, 350	286, 350
93655	Intracardiac catheter ablation of a discrete mechanism of arrhythmia which is distinct from the primary ablated mechanism, including repeat diagnostic maneuvers, to treat a spontaneous or induced arrhythmia (List separately in addition to code for primary procedure)	286, 350	286, 350
93656	Comprehensive electrophysiologic evaluation including transseptal catheterizations, insertion and repositioning of multiple electrode catheters with induction or attempted induction of an arrhythmia with atrial recording and pacing, when possible, right ventricular pacing and recording, His bundle recording with intracardiac catheter ablation of arrhythmogenic focus, with treatment of atrial fibrillation by ablation by pulmonary vein isolation	286, 350	286, 350
93657	Additional linear or focal intracardiac catheter ablation of the left or right atrium for treatment of atrial fibrillation remaining after completion of pulmonary vein isolation	350	350

HERC staff recommendations:

- 1) Remove procedures used solely for ablation of atrial fibrillation from line 286 LIFE-THREATENING CARDIAC ARRHYTHMIAS
 - a. CPT 33254-33259, 33265, 33266
- 2) Remove 33261 (Operative ablation of ventricular arrhythmogenic focus with cardiopulmonary bypass) from lines 73 ACUTE AND SUBACUTE ISCHEMIC HEART DISEASE, MYOCARDIAL INFARCTION and 193 CHRONIC ISCHEMIC HEART DISEASE
 - a. Arrhythmias being treated by this procedure will be on lines 286 or 350
- 3) Add the following guideline to line 350:

GUIDELINE NOTE XXX ABLATION PROCEDURES FOR ATRIAL FIBRILLATION

Line 350

AV nodal ablation (CPT 33250, 33251, 33261, 93650) pairs with atrial fibrillation (ICD-9 427.31/ICD-10 I48.0, I48.1, I48.2, I48.91) only for patients with inadequate ventricular rate control resulting in symptoms, left ventricular systolic dysfunction or substantial risk of left ventricular systolic dysfunction, when pharmacological therapy for rate control is ineffective or not tolerated

Transcatheter pulmonary vein isolation (93656-93657) pairs with atrial fibrillation (ICD-9 427.31/ICD-10 I48.0, I48.1, I48.2, I48.91) only for patients who remain symptomatic from atrial fibrillation despite rate control medications and antiarrhythmic medications.

Surgical ablation (pulmonary vein isolation or Maze procedure) (CPT 33254-33259, 33265, 33266) only pairs with atrial fibrillation (ICD-9 427.31/ICD-10 I48.0, I48.1, I48.2, I48.91) at the time of other cardiac surgery for patients who remain symptomatic despite rate control medications.

HEALTH EVIDENCE REVIEW COMMISSION (HERC)
COVERAGE GUIDANCE: ABLATION FOR ATRIAL FIBRILLATION

DRAFT for VbBS meeting materials 11/13/2014

HERC COVERAGE GUIDANCE

AV node ablation is recommended for coverage only in persons with inadequate ventricular rate control resulting in symptoms, left ventricular systolic dysfunction or substantial risk of left ventricular systolic dysfunction. Coverage is recommended only when pharmacological therapy for rate control is ineffective or not tolerated (*weak recommendation*)

Transcatheter pulmonary vein isolation is recommended for coverage for those who remain symptomatic from atrial fibrillation despite rate control medications and antiarrhythmic medications (*strong recommendation*)

Surgical ablation (pulmonary vein isolation or Maze procedure) for atrial fibrillation is recommended for coverage at the time of other cardiac surgery for patients who remain symptomatic despite rate control medications (*weak recommendation*).

Note: Definitions for strength of recommendation are provided in Appendix A GRADE Element Description

RATIONALE FOR GUIDANCE DEVELOPMENT

The HERC selects topics for guideline development or technology assessment based on the following principles:

- Represents a significant burden of disease
- Represents important uncertainty with regard to efficacy or harms
- Represents important variation or controversy in clinical care
- Represents high costs, significant economic impact
- Topic is of high public interest

Coverage guidance development follows to translate the evidence review to a policy decision. Coverage guidance may be based on an evidence-based guideline developed by the Evidence-based Guideline Subcommittee or a health technology assessment developed by the Health Technology Assessment Subcommittee. In addition, coverage guidance may utilize an existing evidence report produced by one of HERC's trusted sources, generally within the last three years.

EVIDENCE SOURCES

Al-Khatib, S.M., Allen Lapointe, N., Chatterjee, R., Crowley, M.J., Dupre, M.E., Kong, D.F., et al. (2013). *Treatment of atrial fibrillation. Comparative Effectiveness*

Review 119. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2007-10066-I.) AHRQ Publication No.13-EHC095-EF. Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from www.effectivehealthcare.ahrq.gov/reports/final.cfm

Chen, H.S., Wen, J.M., Wu, S.N., & Liu, J.P. (2012). Catheter ablation for paroxysmal and persistent atrial fibrillation. *Cochrane Database of Systematic Reviews, Issue 4*. Art. No.: CD007101. DOI: 10.1002/14651858.CD007101.pub2. Retrieved from

Hashimoto, R.E., Raich, A., Junge, M., & Skelly, A. (2013). Catheter ablation procedures for supraventricular tachyarrhythmia, including atrial flutter & atrial fibrillation. Olympia, WA: Washington State Health Care Authority Health Technology Assessment Program. Retrieved from http://www.hca.wa.gov/hta/Pages/Forms/HTA_Findings.aspx

January, C.T., Wann, L.S., Alpert, J.S., Calkins, H., Cleveland, Jr, J.C., Cigarroa, J.E., et al. (2014). 2014 AHA/ACC/HRS Guideline for the Management of Patients With Atrial Fibrillation: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines and the Heart Rhythm Society. *Circulation*, 129. doi:10.1161/CIR.0000000000000041. Retrieved from <http://circ.ahajournals.org/content/early/2014/03/27/CIR.0000000000000041.citation>

The summary of evidence in this document is derived directly from these evidence sources, and portions are extracted verbatim.

SUMMARY OF EVIDENCE

Clinical Background

Atrial fibrillation (AF) is a supraventricular tachyarrhythmia characterized by uncoordinated atrial activation with consequent deterioration of mechanical function. Different systems have been proposed to classify AF. Although the type of AF can change over time, it is often helpful to characterize it at a given moment, as this may guide treatment. Types of AF include first-detected, paroxysmal (arrhythmia terminates spontaneously within 7 days), persistent (arrhythmia is sustained beyond 7 days), longstanding persistent (patients who have been in AF for any period longer than 1 year when attempts at achieving sinus rhythm are planned or are in progress), and permanent AF (in which cardioversion has failed or has not been attempted).

It is estimated that more than 2.3 million Americans have AF. The prevalence of AF increases with age and approaches 8 percent in patients older than 80 years of age. AF is the most common sustained arrhythmia seen in clinical practice. The impact of AF is

compounded by its known association with significant mortality, morbidity, and health care costs. Not only is the risk of death in patients with AF twice that of patients without AF, but AF can result in myocardial ischemia or even infarction, heart failure exacerbation, and tachycardia-induced cardiomyopathy if the ventricular rate is not well controlled. The most dreaded complication of AF is thromboembolism, especially stroke. Importantly, when ischemic stroke occurs in patients with AF, it is either fatal or of moderate to high severity in the majority of patients. The management of AF and its complications is responsible for almost \$16 billion in costs to the U.S. health care system each year.

Treatment Strategies

Management of AF involves three distinct areas: rate control (treatments to slow the heart rate to a normal range), rhythm control (treatments to revert the heart rhythm back to normal), and prevention of thromboembolic events. Whether or not a rhythm-control strategy is adopted, current treatment guidelines suggest that adequate rate control should be achieved in all patients with AF to prevent myocardial infarction (if significant coronary artery disease is present), exacerbation of heart failure, and tachycardia-induced cardiomyopathy; to alleviate symptoms; and to improve exercise tolerance and quality of life.

Rate Control

If pharmacological therapy is insufficient for rate control and symptom management or is associated with side effects, the 2006 ACC/AHA/ESC Guidelines recommend ablation of the atrioventricular node (AVN) in conjunction with permanent pacemaker implantation to control heart rate. As the latter involves implantation of an indwelling device that is not reversible, it is considered a treatment of last resort for patients for whom initial pharmacotherapy was ineffective.

Another clinical dilemma is whether patients with AF do better with strict or lenient rate control. In theory, strict control could reduce symptoms and prevent complications. However, stricter control requires more intensive use of medications, which carry their own side effects. The 2011 Focused Update on the Management of Patients With Atrial Fibrillation by the American College of Cardiology Foundation (ACCF), the AHA, and the Heart Rhythm Society (HRS) addressed the issue of strict versus lenient rate control in patients with AF. Specifically, these guidelines emphasized the following Class III recommendation (evidence and/or general agreement that the procedure/treatment is not useful/effective and in some cases may be harmful): "Treatment to achieve strict rate control of heart rate (<80 bpm at rest or <110 bpm during a 6-minute walk) is not beneficial compared with achieving a resting heart rate <110 bpm in patients with persistent AF who have stable ventricular function (left ventricular ejection fraction >0.40) and no or acceptable symptoms related to the arrhythmia."

Rhythm Control

If patients with AF continue to have significant symptoms despite adequate rate control through either pharmacological therapy or AVN ablation, then a rhythm-control strategy (either pharmacological or electrical) is currently recommended. For pharmacological cardioversion of AF, the 2014 ACC/AHA/ESC Guidelines recommend flecainide, dofetilide, propafenone, and ibutilide as Class I recommendations, and amiodarone as a Class IIa recommendation (weight of evidence/opinion is in favor of usefulness/efficacy). To enhance direct-current cardioversion, the 2014 ACC/AHA/ESC Guidelines recommend pretreatment with amiodarone, flecainide, ibutilide, propafenone, or sotalol. For maintenance of sinus rhythm after cardioversion, the 2014 ACC/AHA/ESC Guidelines list different antiarrhythmic medications for different clinical settings.

In addition to pharmacological and direct-current cardioversion, a number of surgical interventions are used for rhythm control. Catheter ablation for the treatment of AF, with pulmonary vein isolation (PVI) being the most commonly used ablation, has evolved rapidly from a highly experimental procedure to its current status as a commonly performed procedure that is widely regarded as a clinically useful treatment option for symptomatic patients with AF in whom medications are not effective or not tolerated.

Several other procedures for the treatment of AF have been investigated. One such procedure is the surgical Maze procedure, which appears to confer some benefit to selected patients with AF. Although several studies of rate- and rhythm-control strategies exist, to date no study has shown that maintaining patients with AF in sinus rhythm provides a long-term survival benefit. It is also unknown whether the risks and benefits of different therapies vary by AF type.

Evidence Review

Rate-Control Procedures Versus Drugs or Versus Other Procedures in Patients for Whom Initial Pharmacotherapy Was Ineffective

Al-Khatib 2013 reports on four RCTs (one good, two fair, and one poor quality) involving a total of 211 patients that compared the effectiveness of a procedural intervention versus a primarily pharmacological intervention for rate control of AF. All four studies recruited patients with permanent AF, (referred to as “resistant chronic” AF in one study). All studies included at least one treatment arm with radiofrequency ablation of either the AVN or His bundle in conjunction with pacemaker placement. The comparison arms included a pharmacological intervention whose main purpose was to control ventricular heart rate rather than converting the underlying rhythm of AF.

Based on three studies reported in Al-Khatib 2013 (one good, one fair, one poor quality) involving 211 patients, patients undergoing a procedural intervention had a significantly lower heart rate at 12 months than those receiving a primarily pharmacological intervention. This was measured differently in all three studies. In one, the mean heart rate in the intervention group was 71 ± 6 bpm compared to 83 ± 8 bpm in the medication group ($p < 0.01$). In this study, maximum heart rate did not differ between groups. In the second study, those in the ablation group had higher minimum (70 ± 9 vs. 39 ± 9 bpm; $p < 0.05$) and mean (76 ± 7 vs. 71 ± 11 bpm; $p < 0.05$) heart rates than the medication group, but lower maximum heart rates (117 ± 16 bpm vs. 152 ± 37 bpm; $p < 0.05$). The third study reported the percent of each group who had either a normal or uncontrolled ventricular rate; in the ablation group, 100% had a normal ventricular rate (50-90 bpm) compared to 58% in the medication group. Similarly, none of the ablation group had an uncontrolled heart rate (> 90 bpm at rest or > 130 bpm on exertion), while 42% of the medication group did. There was no difference by treatment arm in all-cause mortality (two studies [one good, one fair quality], 201 patients); cardiovascular mortality (one study [good quality], 102 patients); or exercise capacity (two studies [one good, one fair quality], 135 patients) (all low strength of evidence). There was insufficient strength of evidence to support findings for other outcomes, including quality of life.

Rhythm-Control Procedures and Drugs for Maintenance of Sinus Rhythm

Al-Khatib 2013 included 65 studies enrolling 6,739 patients that evaluated procedures for rhythm control. Of those that specified type of AF, eleven included only patients with longstanding persistent AF, 17 studies included only patients with paroxysmal AF, and 4 studies included only patients with persistent AF.

Transcatheter PVI versus antiarrhythmic drugs

Al-Khatib 2013 concluded, based on eight RCTs (five good, three fair quality) involving 921 patients, that transcatheter PVI is superior to antiarrhythmic drugs for **maintenance of sinus rhythm** over 12 months of follow up in patients with AF (one RCT reported 48 months of follow up). All trials had statistically significant results, as did meta-analysis of all eight trials (OR 6.51, 95% CI 3.22 to 13.16). This evidence is strongest in younger patients with little to no structural heart disease and with mild or no enlargement of the left atrium. Only one trial was limited to patients receiving ablation as first line therapy (Wazni 2005), while five specifically required failure of at least one AAD to be included in the study. The Wazni trial included 70 patients who experienced monthly episodes of symptomatic AF for at least three months, and found that at one year follow up, 63% of those treated with AADs had at least one recurrence of AF, compared to 13% of those

who received PVI. Another trial included only in Hashimoto 2013 included only patients with persistent AF (MacDonald 2011), and reported that at final follow up (6 months), 50% of patients in the PVI group were in sinus rhythm while none of the control group were (no statistical testing done). This latter trial was limited to patients with advanced heart failure. (Note: This outcome is reported as freedom from recurrence in Hashimoto 2013, but results are similar.)

Al-Khatib 2013 concluded, based on two RCTs (Pappone 2006, Forleo 2009, both good quality) involving 268 patients, that transcatheter PVI is superior to antiarrhythmic medications in **reducing cardiovascular hospitalizations** (moderate strength of evidence). Both of these trials were also included in Hashimoto 2013. A third study, Stabile 2006, reported only in Hashimoto 2013, found a lower number of hospitalizations in the PVI group which did not reach statistical significance. A fourth RCT, Wazni 2005, reported only in Al-Khatib 2013, found the rate of hospitalization specifically for AF was higher in the AAD arm (15 of 35) than the PVI arm (3 of 32, $p < 0.001$) in the first 12 months of follow up.

Chen 2012 reported that only one trial (Stabile 2006) reported all-cause **mortality**. There were no statistically significant differences between groups for this outcome. In this trial, the one death that occurred in the PVI group was from a stroke that occurred during the procedure and was followed by a brain hemorrhage 9 months later. There were two deaths in the AAD group (diagnosis not specified).

Al-Khatib 2013 also reported only one study for the outcome of all-cause mortality, however, it was a different study than was reported by Chen. This study (Oral 2006) reported one death in the PVI arm at 12 months compared to none in the AAD arm; no statistical testing was done.

Hashimoto 2013 reported that four RCTs (Jais 2008, Wilbur 2010, Stabile 2006, Oral 2006) reported overall mortality rates (not procedure related) at 9 to 12 months of follow up. Mortality rate in the PVI arm ranged from 1% to 3%, while in the AAD arm a rate of 3% was reported in two studies. According to Hashimoto, Stabile 2006 was the only RCT to report mortality in both arms. Two cohort studies included in Hashimoto 2013 did report an increased risk of death in the AAD group at follow up times ranging from 1 to 3 years (Pappone 2003: 6.5% in the PVI group vs. 14.3% in the AAD group, $p < 0.001$) or at a mean follow up of 69 months (Sonne 2009: 2.1% in the PVI group vs. 16.5% in the AAD group, $p = 0.001$).

Eight studies evaluated **quality of life (QOL) or functional status**, three RCTs reported in all three source reports, two additional RCTs reported in both Hashimoto 2013 and Al-Khatib 2013, two additional RCTs in Hashimoto 2013 only and one cohort study reported in Al-Khatib 2013 only. In general, there was greater improvement from

baseline in these scores in patients randomized to the PVI arm, compared to the AAD arm, and in most of these studies, results were statistically significant for at least some measures.

Harms were reported in eight RCTs, but for the most part, were not statistically analyzed. Complications reported in each study are summarized in the Table below:

Author	N	PVI Arm	AAD Arm
Krittayaphong	30	1 stroke, 1 groin hematoma	AE in 7 patients (47%): GI AE in 6 pts, corneal deposits in 2 pts, hypothyroidism in 2 pts, abnormal LFTs in 2 pts, hyperthyroidism in 1 pt, sinus node dysfunction in 1 pt
Wazni	70	No TE events, no bradycardia, 1 asymptomatic PV stenosis	No TE events, 8.6% bradycardia
Pappone	198	No serious AE	Sig AE leading to drug withdrawal in 23 pts,
Oral	146	None	None
Stabile	137	4.4% major complications (stroke, phrenic paralysis, pericardial effusion)	1 TIA, 2 cancer, 1 sudden death
Jais	112	2 cardiac tamponade, 2 groin hematomas, 1 PV stenosis requiring stent	1 hyperthyroidism, 2 deaths (unrelated)
Forleo	70	1 groin hematoma	17% sig drug AE (bradycardia, atrial flutter, sinus node dysfunction)
Wilber	167	5 major AE (pericardial effusion, pulmonary edema, pneumonia, vascular complication, heart failure)	5 major AE (2 life-threatening arrhythmias, 3 disabling drug intolerance requiring discontinuation)

TE = thromboembolic; PV = pulmonary vein

Cryoablation PVI vs. AAD

One RCT reported in Hashimoto 2013 found that patients randomized to receive cryoablation had significantly greater freedom from recurrence compared with those patients randomized to receive AADs alone (69.9% versus 7%, respectively; $P < .001$). There was one death (0.6%) in the cryoablation PVI group and none in the AAD group at 12 months, which was not statistically significant.

Surgical Maze versus standard of care (mitral valve surgery)

Al-Khatib 2013 included seven RCTs (one good, six fair quality) involving 361 patients for this comparison. Surgical Maze at the time of other cardiac surgery (specifically mitral valve surgery) is superior to mitral valve surgery alone for **maintenance of sinus rhythm** over at least 12 months of followup in patients with persistent AF (OR 5.80, 95% CI 1.79 to 18.81). Six studies reported on all cause mortality; meta-analysis found an OR of 1.97 (95% CI 0.81 to 4.80) suggesting an increased risk of death with the Maze procedure, but this did not reach statistical significance.

PVI done at the time of cardiac surgery versus cardiac surgery alone or cardiac surgery in combination with antiarrhythmic drugs (AADs) or catheter ablation

Al-Khatib 2013 included eight RCTs (five good, three fair quality) involving 532 patients for this comparison. Pulmonary vein isolation done at the time of cardiac surgery is superior to cardiac surgery alone or cardiac surgery in combination with AADs or catheter ablation for **maintenance of sinus rhythm** over 12 months of followup in patients with persistent AF (OR 3.91, 95% CI 1.54 to 9.91). Two studies reported no difference between groups in all-cause mortality or stroke.

There are insufficient data on the effect of rhythm control with PVI or surgical Maze on final outcomes, such as all-cause mortality, stroke, heart failure, and left ventricular ejection fraction, and on the safety and durability of the effectiveness of these procedures beyond 12 months.

Other comparisons

There are a variety of other comparisons included in Al-Khatib 2013 and Chen 2012, most of which had a limited number of studies and were considered outside the scope of this guidance document. These include the following:

- Circumferential PVI versus Segmental PVI
- Transcatheter PVI with complex fractionated atrial electrogram (CFAE) ablation versus transcatheter PVI without CFAE ablation
- Transcatheter PVI using different types of ablation catheters
- Transcatheter PVI with Cavotricuspid isthmus (CTI) ablation vs. transcatheter PVI without CTI ablation
- Transcatheter PVI vs transcatheter PVI with ablation sites other than CTI and CFAE and transcatheter PVI involving all four PVs vs transcatheter PVI involving arrhythmogenic PVs only

- Transcatheter PVI Alone vs transcatheter PVI plus postablation antiarrhythmic drugs
- Left atrial ablation vs. bi-atrial ablation
- PVI, circumferential PVI or left atrium ablation vs. ablation plus additional linear ablation
- PV-left atrium junction ablation vs. PV-left atrium junction ablation combined with CTI ablation
- Circumferential PV ablation vs. circumferential PV ablation plus PVI
- Superior PV ablation vs. four-PV ablation
- Small area isolation vs. large area isolation around PVs in circumferential PV ablation
- CFAE plus PV atrium isolation vs. PV atrium isolation alone
- Circumferential PV ablation vs. modified circumferential PV ablation
- Arrhythmogenic PVI vs all PVI

[Evidence Source]

Evidence Summary

Ablation of the AV node or bundle of His in patients with AF results in lower heart rate at 12 months than pharmacologic treatment (moderate SOE), although there is no difference in mortality or exercise capacity (low SOE). Pulmonary vein isolation (PVI) results in a greater likelihood of maintaining sinus rhythm at 12 months than pharmacologic treatment (high SOE); most of the evidence for this finding is in patients with AF who have failed at least one AAD. This procedure (PVI) also results in lower risk of hospitalization over 12 months (moderate SOE) and improved QOL (moderate SOE), but the evidence is insufficient to assess the impact of PVI on mortality.

The surgical Maze procedure, when done at the time of other cardiac surgery, results in a higher likelihood of maintaining sinus rhythm than not performing the Maze (moderate SOE). Similarly, PVI done at the time of other cardiac surgery results in a higher likelihood of maintaining sinus rhythm than not performing PVI (high SOE), and no apparent difference in all-cause mortality or stroke (low SOE).

GRADE-INFORMED FRAMEWORK

The HERC develops recommendations by using the concepts of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system. GRADE is a transparent and structured process for developing and presenting evidence and for carrying out the steps involved in developing recommendations. There are four elements that determine the strength of a recommendation, as listed in the table below. The HERC reviews the evidence and makes an assessment of each element, which in turn is used to develop the recommendations presented in the coverage guidance box. Balance between desirable and undesirable effects, and quality of evidence, are derived from the evidence presented in this document, while estimated relative costs, values and preferences are assessments of the HERC members.

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variability in values and preferences	Coverage recommendation	Rationale
Ablation of AV node/bundle of His compared to rate control medications in patients for whom initial pharmacotherapy was ineffective	Lower heart rate, no difference in mortality/exercise capacity	Moderate/Low based on 1 to 3 poor to good quality studies, depending on the outcome	High	High	AV node ablation is recommended for coverage only in symptomatic persons when pharmacological therapy for rate control is ineffective or not tolerated. (<i>weak recommendation</i>)	Studies show mixed clinical significance of a lower heart rate. In those with persistently uncontrolled heart rate despite AADs, AV node ablation is a reasonable alternative to prevent the negative consequences of an uncontrolled rate such as MI, exacerbation of CHF or cardiomyopathy.
Transcatheter PVI vs. AAD	Better maintenance of SR, fewer	High/Moderate, based on 1 to	High	Moderate	Transcatheter PVI is recommended for	Transcatheter PVI produces

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variability in values and preferences	Coverage recommendation	Rationale
	hospitalizations, better QOL, possibly lower mortality	8 fair to good quality studies, depending on the outcome			coverage when a rhythm control strategy is desired (<i>strong recommendation</i>)	superior clinical outcomes to antiarrhythmic drugs alone when a rhythm control strategy is pursued
Maze procedure	Better maintenance of SR; possible (nonsignificant) increase in mortality	Moderate based on 1 good and six fair quality studies	Moderate (concurrent with other cardiac surgery)	Moderate	The Maze procedure is recommended for coverage at the time of other cardiac surgery if the benefits of maintenance of sinus rhythm are thought to outweigh the potential risk of increased mortality (<i>weak recommendation</i>)	Maze may help maintain sinus rhythm but concerning nonsignificant increased risk of mortality
PVI done with other cardiac surgery	Better maintenance of SR	High based on 5 good and 3 fair quality studies	Moderate (concurrent with other cardiac surgery)	Low	PVI is recommended for coverage (<i>weak recommendation</i>)	PVI may help maintain sinus rhythm without significant additional risks

SR = sinus rhythm PVI = pulmonary vein isolation AAD = anti-arrhythmic drugs

*The Quality of Evidence rating was assigned by the primary evidence source, not the HERC Subcommittee

Note: GRADE framework elements are described in Appendix A

POLICY LANDSCAPE

Nine quality measures pertaining to atrial fibrillation were identified when searching the [National Quality Measures Clearinghouse](#); however, none of them referenced ablation.

Choosing Wisely[®] is part of a multi-year effort of the ABIM Foundation to help physicians be better stewards of finite health care resources. Originally conceived and piloted by the National Physicians Alliance through a Putting the Charter into Practice grant, more than 50 medical specialty organizations, along with Consumer Reports, have identified a number of tests or procedures commonly used in their field, whose necessity should be questioned and discussed. Each participating organization was free to determine how to create its own list, provided that it used a clear methodology and adhered to the following set of shared guidelines:

- Each item should be within the specialty's purview and control.
- The tests and/or interventions should be used frequently and/or carry a significant cost.
- Each recommendation should be supported by generally accepted evidence.
- The selection process should be thoroughly documented and publicly available on request.

One of the organizations that chose to participate in the *Choosing Wisely*[®] campaign is the Heart Rhythm Society. The most recent list created by this organization states the following:

“Don’t ablate the atrioventricular node in patients with atrial fibrillation when both symptoms and heart rate are acceptably controlled by well-tolerated medical therapy.

Atrioventricular node ablation and pacemaker implantation may provide benefit in some patients when rate and related symptoms cannot be controlled by medication therapy,(Class IIa, indicated) or when there is concern for possible tachycardia-induced cardiomyopathy (Class IIb, may be considered). However, according to current professional society clinical guidelines, the risks of AV node ablation outweigh the benefits among patients with no symptoms and who have appropriate rate control with well-tolerated medical therapy.”

They cite the 2011 publication of the ACCF/AGA guidelines on the management of patient with AF as supporting evidence. These guidelines were recently updated (2014), and are rated fair quality using the MED standard criteria, primarily because study selection criteria was not specified and the quality of included studies was not

assessed. These guidelines state the following with regard to AV node ablation for rate control in AF:

Class IIa

3. AV nodal ablation with permanent ventricular pacing is reasonable to control the heart rate when pharmacological therapy is inadequate and rhythm control is not achievable. (Level of Evidence: B)

Class III: Harm

1. AV nodal ablation with permanent ventricular pacing should not be performed to improve rate control without prior attempts to achieve rate control with medications. (Level of Evidence: C)

For catheter ablation for rhythm control (e.g. PVI), the guidelines state the following:

Class I

1. AF catheter ablation is useful for symptomatic paroxysmal AF refractory or intolerant to at least 1 class I or III antiarrhythmic medication when a rhythm control strategy is desired. (Level of Evidence: A)

2. Prior to consideration of AF catheter ablation, assessment of the procedural risks and outcomes relevant to the individual patient is recommended. (Level of Evidence: C)

Class IIa

1. AF catheter ablation is reasonable for selected patients with symptomatic persistent AF refractory or intolerant to at least 1 class I or III antiarrhythmic medication. (Level of Evidence: A)

2. In patients with recurrent symptomatic paroxysmal AF, catheter ablation is a reasonable initial rhythm control strategy prior to therapeutic trials of antiarrhythmic drug therapy, after weighing risks and outcomes of drug and ablation therapy. (Level of Evidence: B)

Class IIb

1. AF catheter ablation may be considered for symptomatic long-standing (>12 months) persistent AF refractory or intolerant to at least 1 class I or III antiarrhythmic medication, when a rhythm control strategy is desired. (Level of Evidence: B)

2. AF catheter ablation may be considered prior to initiation of antiarrhythmic drug therapy with a class I or III antiarrhythmic medication for symptomatic persistent AF, when a rhythm control strategy is desired. (Level of Evidence: C)

Class III: Harm

1. AF catheter ablation should not be performed in patients who cannot be treated with anticoagulant therapy during and following the procedure. (*Level of Evidence: C*)

2. AF catheter ablation to restore sinus rhythm should not be performed with the sole intent of obviating the need for anticoagulation. (*Level of Evidence: C*)

Coverage guidance is prepared by the Health Evidence Review Commission (HERC), HERC staff, and subcommittee members. The evidence summary is prepared by the Center for Evidence-based Policy at Oregon Health & Science University (the Center). This document is intended to guide public and private purchasers in Oregon in making informed decisions about health care services.

The Center is not engaged in rendering any clinical, legal, business or other professional advice. The statements in this document do not represent official policy positions of the Center. Researchers involved in preparing this document have no affiliations or financial involvement that conflict with material presented in this document.

Appendix A. GRADE Element Descriptions

Element	Description
Balance between desirable and undesirable effects	The larger the difference between the desirable and undesirable effects, the higher the likelihood that a strong recommendation is warranted. The narrower the gradient, the higher the likelihood that a weak recommendation is warranted
Quality of evidence	The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted
Resource allocation	The higher the costs of an intervention—that is, the greater the resources consumed—the lower the likelihood that a strong recommendation is warranted
Values and preferences	The more values and preferences vary, or the greater the uncertainty in values and preferences, the higher the likelihood that a weak recommendation is warranted

Strong recommendation

In Favor: The subcommittee is confident that the desirable effects of adherence to a recommendation outweigh the undesirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences.

Against: The subcommittee is confident that the undesirable effects of adherence to a recommendation outweigh the desirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences.

Weak recommendation

In Favor: The subcommittee concludes that the desirable effects of adherence to a recommendation probably outweigh the undesirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences, but is not confident.

Against: The subcommittee concludes that the undesirable effects of adherence to a recommendation probably outweigh the desirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences, but is not confident.

Quality or strength of evidence rating across studies for the treatment/outcome¹

High: The subcommittee is very confident that the true effect lies close to that of the estimate of the effect. Typical sets of studies are RCTs with few or no limitations and the estimate of effect is likely stable.

Moderate: The subcommittee is moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Typical sets of studies are RCTs with some limitations or well-performed nonrandomized studies with additional strengths that guard against potential bias and have large estimates of effects.

Low: The subcommittee's confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. Typical sets of studies are RCTs with serious limitations or nonrandomized studies without special strengths.

Very low: The subcommittee has very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect. Typical sets of studies are nonrandomized studies with serious limitations or inconsistent results across studies.

¹ Includes risk of bias, precision, directness, consistency and publication bias

Appendix B. Applicable Codes

CODES DESCRIPTION	
ICD-9 Diagnosis Codes	
427.31	Atrial fibrillation
ICD-10 Diagnosis Codes	
I48.0	Paroxysmal atrial fibrillation
I48.1	Persistent atrial fibrillation
I48.2	Chronic atrial fibrillation
I48.91	Unspecified atrial fibrillation
ICD-9 Volume 3 (Procedure Codes)	
None	
CPT Codes	
33250	Operative ablation of supraventricular arrhythmogenic focus or pathway (eg, Wolff-Parkinson-White, atrioventricular node re-entry), tract(s) and/or focus (foci); without cardiopulmonary bypass (For intraoperative pacing and mapping by a separate provider, use 93631) Codes 33254-33256 are only to be reported when there is no concurrently performed procedure that requires median sternotomy or cardiopulmonary bypass.
33251	...with cardiopulmonary bypass
33254	Operative tissue ablation and reconstruction of atria, limited (eg, modified maze procedure)
33255	Operative tissue ablation and reconstruction of atria, extensive (eg, maze procedure); without cardiopulmonary bypass
33256	...with cardiopulmonary bypass
33257	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), limited (eg, modified maze procedure) (List separately in addition to code for primary procedure)
33258	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), extensive (eg, maze procedure), without cardiopulmonary bypass (List separately in addition to code for primary procedure)
33259	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), extensive (eg, maze procedure), with cardiopulmonary bypass (List separately in addition to code for primary procedure)
33261	Operative ablation of ventricular arrhythmogenic focus with cardiopulmonary bypass
33265	Endoscopy, surgical; operative tissue ablation and reconstruction of atria, limited (eg, modified maze procedure), without cardiopulmonary bypass
33266	...operative tissue ablation and reconstruction of atria, extensive (eg, modified maze procedure), without cardiopulmonary bypass
93613	Intracardiac electrophysiologic 3-dimensional mapping (List separately in addition to code for primary procedure)
93650	Intracardiac catheter ablation of atrioventricular node function, atrioventricular conduction for creation of complete heart block, with or without temporary pacemaker placement
93653	Comprehensive electrophysiologic evaluation including insertion and repositioning of multiple electrode catheters with induction or attempted induction of an arrhythmia with right atrial pacing and recording, right ventricular pacing and recording, His recording with intracardiac catheter ablation of arrhythmogenic focus; with treatment

CODES	DESCRIPTION
	of supraventricular tachycardia by ablation of fast or slow atrioventricular pathway, accessory atrioventricular connection, cavo-tricuspid isthmus or other single atrial focus or source of atrial re-entry (Do not report 93653 in conjunction with 93600-93603, 93610, 93612, 93618-93620, 93642, 93654)
93655	Intracardiac catheter ablation of a discrete mechanism of arrhythmia which is distinct from the primary ablated mechanism, including repeat diagnostic maneuvers, to treat a spontaneous or induced arrhythmia (List separately in addition to code for primary procedure) (Use 93655 in conjunction with 93653, 93654, 93656)
93656	Comprehensive electrophysiologic evaluation including transseptal catheterizations, insertion and repositioning of multiple electrode catheters with induction or attempted induction of an arrhythmia with atrial recording and pacing, when possible, right ventricular pacing and recording, His bundle recording with intracardiac catheter ablation of arrhythmogenic focus, with treatment of atrial fibrillation by ablation by pulmonary vein isolation
93657	Additional linear or focal intracardiac catheter ablation of the left or right atrium for treatment of atrial fibrillation remaining after completion of pulmonary vein isolation (List separately in addition to code for primary procedure)
93799	Unlisted cardiovascular service or procedure
HCPSC Level II Codes	
None	

Note: Inclusion on this list does not guarantee coverage

Appendix C. HERC Guidance Development Framework

HERC Guidance Development Framework Principles

This framework was developed to assist with the decision making process for the Oregon policy-making body, the HERC and its subcommittees. It is a general guide, and must be used in the context of clinical judgment. It is not possible to include all possible scenarios and factors that may influence a policy decision in a graphic format. While this framework provides a general structure, factors that may influence decisions that are not captured on the framework include but are not limited to the following:

- Estimate of the level of risk associated with the treatment, or any alternatives;
- Which alternatives the treatment should most appropriately be compared to;
- Whether there is a discrete and clear diagnosis;
- The definition of clinical significance for a particular treatment, and the expected margin of benefit compared to alternatives;
- The relative balance of benefit compared to harm;
- The degree of benefit compared to cost; e.g., if the benefit is small and the cost is large, the committee may make a decision different than the algorithm suggests;
- Specific indications and contraindications that may determine appropriateness;
- Expected values and preferences of patients.

Ablation of AV node/bundle of His vs. rate control medications



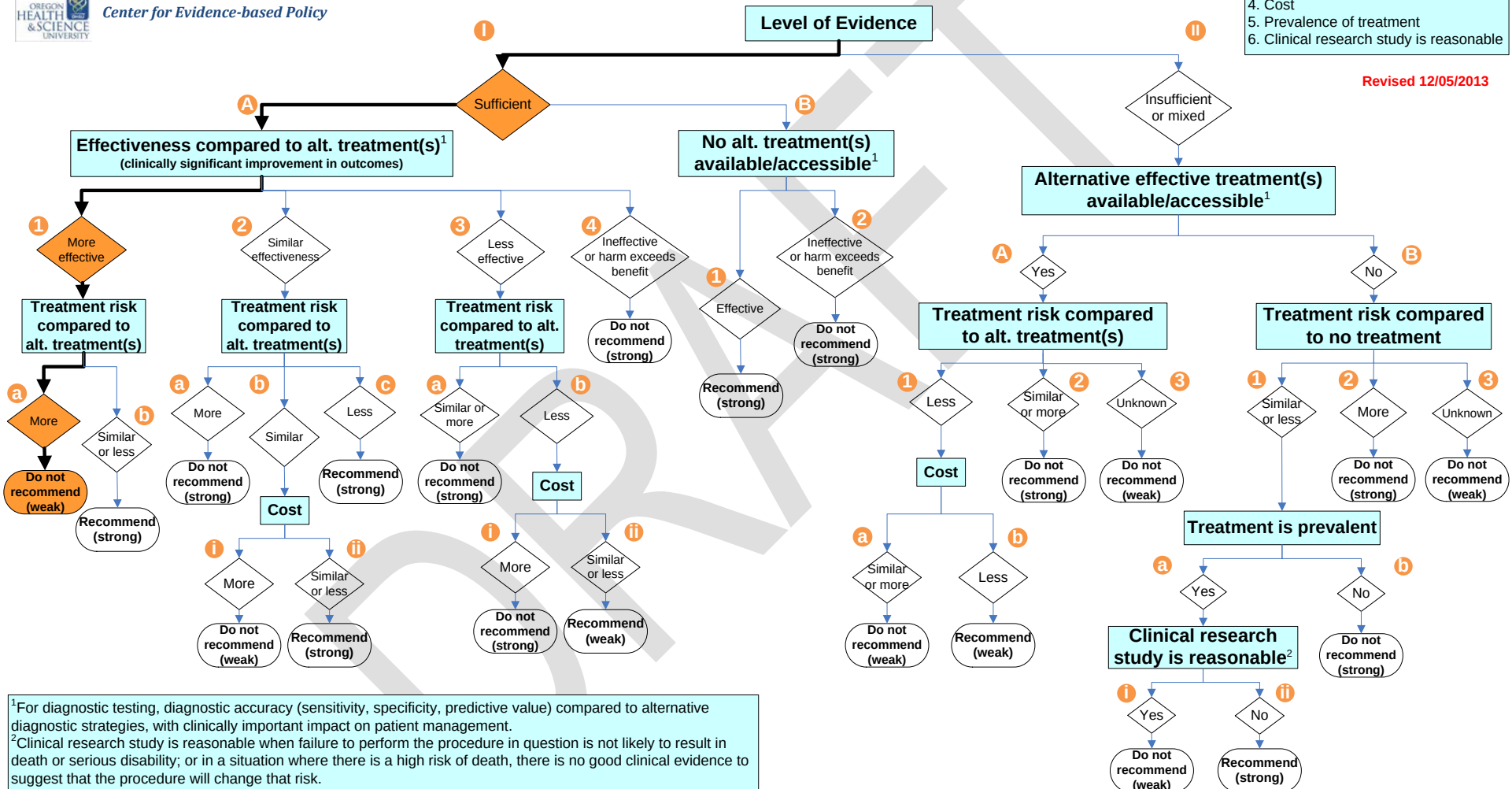
Center for Evidence-based Policy

HERC Guidance Development Framework

Refer to HERC Guidance Development Framework Principles for additional considerations

- Decision Point Priorities**
1. Level of evidence
 2. Effectiveness & alternative treatments
 3. Harms and risk
 4. Cost
 5. Prevalence of treatment
 6. Clinical research study is reasonable

Revised 12/05/2013



Transcatheter pulmonary vein isolation (PVI) vs. antiarrhythmic drugs (AAD); Surgical ablation (Maze procedure or PVI done with other cardiac surgery)



Center for Evidence-based Policy

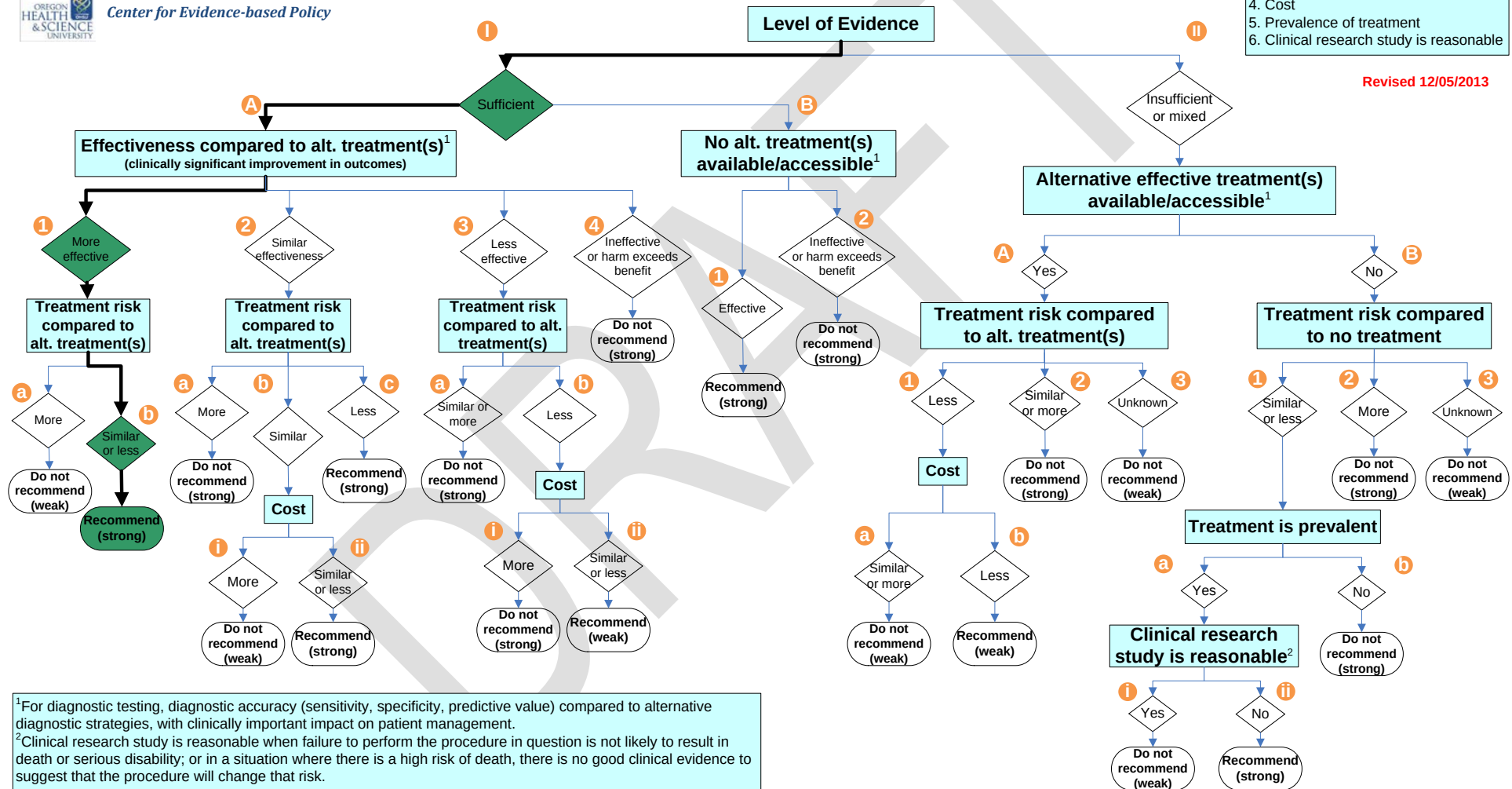
HERC Guidance Development Framework

Refer to *HERC Guidance Development Framework Principles* for additional considerations

Decision Point Priorities

1. Level of evidence
2. Effectiveness & alternative treatments
3. Harms and risk
4. Cost
5. Prevalence of treatment
6. Clinical research study is reasonable

Revised 12/05/2013



HERC Coverage Guidance – Ablation for Atrial Fibrillation Disposition of Public Comments

Table of Contents

Commenters.....	1
Public Comments	2
References Provided by Commenters	6

Commenters

Identification	Stakeholder
A	Family Physician, Reedsport, OR [Submitted June 26, 2014]
B	Oregon Chapter of the American College of Cardiology, Portland, OR [Submitted July 22, 2014]
C	Medtronic, Inc., Mounds View, MN [Submitted July 21, 2014]

HERC Coverage Guidance – Ablation for Atrial Fibrillation Disposition of Public Comments

Public Comments

Ident.	#	Comment	Disposition
A	1	Agree with recommendations. They appear to be well-researched. Choosing Wisely recommendation adds weight; these have been researched by others and are endorsed by the specialty society.	Thank you for your comment.
B	1	The Oregon Chapter of the American College of Cardiology appreciates the opportunity to provide comments on the Health Evidence Review Commission's Coverage Guidance for Ablation for the Treatment of Atrial Fibrillation. We are of the opinion that the coverage guidance is well written and appropriate and agree with the overall conclusions and coverage decisions. However, there are several points worth revising, as detailed below.	Thank you for your comment.
B	2	Page 1: Box labeled "HERC Coverage Guidance", the third and fourth points are redundant, i.e., PVI at time of other cardiac surgery and surgical MAZE procedure at time of other cardiac surgery are the same. This error is repeated several times in the document. Although Al-Khatib distinguishes these two terms in her paper, these are generally used as synonyms in clinical practice, and thus comparisons between them are not valid. We recommend that the definitions of PVI and MAZE be included in the guidance.	PVI and surgical Maze are discussed separately in the AHRQ review, with different studies informing conclusions. Neither procedure is defined or described in this report. Given that both procedures have similar recommendations, box language changed to the generic term "surgical ablation".
B	3	Page 2: First paragraph under "clinical background", the term "permanent afib" is now rarely used and signifies afib that is always present and that no attempts are being made to restore sinus rhythm. We suggest clarifying that if attempts are made, then the afib is referred to as "longstanding persistent".	Current description of atrial fibrillation varieties is from the AHRQ report, and states the following: "Types of AF include first-detected, paroxysmal (arrhythmia terminates spontaneously within 7 days), persistent (arrhythmia is sustained beyond 7 days), longstanding persistent (usually lasting for more than 1 year <u>when attempts at achieving sinus rhythm are planned or are in progress</u>), and permanent AF (in which cardioversion has failed or has not been attempted)." Suggested clarification made to the text.
B	4	Page 3: Top, "the risk of stroke is up to 8%, depending on other stroke risk factors". To more appropriately classify this risk, we recommend the statement be revised to state a range, such as 1-8%.	Current language is verbatim from the AHRQ report. No citation provided for the range. AHA guideline reports 5X

HERC Coverage Guidance – Ablation for Atrial Fibrillation Disposition of Public Comments

Ident.	#	Comment	Disposition
			increase in stroke risk, but does not list an actual %. Sentence deleted.
B	5	Page 4: Middle, cardiac resynchronization therapy (CRT) is presented as a potential treatment for afib. CRT can be used in conjunction with AV node ablation, but the goal is not to decrease afib burden, as is currently discussed. Existing practice guidelines and current clinical practice do not include CRT as an appropriate treatment for afib, and we recommend removing it as a treatment option.	Language is from the background section and is verbatim from the AHRQ report; there is no evidence pertaining to CRT in the evidence review section. Sentence deleted.
B	6	Page 5: Top, AV node ablation is always in conjunction with placement of a permanent pacemaker (or previous placement of such). We suggest revision to reflect this.	This language is verbatim from the AHRQ report; four of the six included studies evaluated AV node ablation compared to pharmacologic treatment; all of these included placement of a pacemaker. One of the remaining studies compared two different approaches (anterior and posterior) to “AV junction modification”. In this study, no pacemaker was used, as AV node pacing function remained intact. The sixth study compared right ventricular pacing to biventricular pacing. Text of document revised to eliminate reference to studies that did not compare a primarily pharmacologic intervention to a primarily procedural intervention; this allows deletion as follows: “All studies included at least one treatment arm with radiofrequency ablation of either the AVN or His bundle, ~~most often~~ in conjunction with pacemaker placement.”
B	7	Page 8: There are separate paragraphs dealing with surgical PVI v. MAZE procedure at time of other cardiac surgery. Again, see page 1 comments above.	See comment #B2. Box language changed to surgical ablation. However,

HERC Coverage Guidance – Ablation for Atrial Fibrillation Disposition of Public Comments

Ident.	#	Comment	Disposition
			because these two procedures are addressed discretely in the AHRQ report, with different studies supporting PVI and Maze, text of the document other than the box not revised.
B	8	Page 9: Bottom, again, surgical PVI and MAZE are the same thing.	See comment #B7
C	1	Medtronic appreciates this opportunity to comment on Oregon Health Evidence Review Commission's (HERC's) draft coverage guidance on ablation for atrial fibrillation (AF). ¹ Medtronic has extensive clinical expertise and offers innovative products across several areas of cardiovascular care, including AF. We applaud the efforts of HERC to develop evidence-based coverage guidance on treatments for AF. The guidance clearly summarizes and evaluates the robust body of evidence for AF treatment strategies.	Thank you for your comment.
C	2	To ensure alignment with the clinical guidelines and avoid potential confusion, we request clarification of the term "persistently symptomatic" in the draft coverage language for pulmonary vein isolation (PVI). The coverage guidance currently states: <i>"Transcatheter pulmonary vein isolation is recommended for coverage for those who are persistently symptomatic despite rate control medications and antiarrhythmic medications (strong recommendation)."</i> This language may suggest coverage is limited to patients with <i>persistent</i> AF and does not include patients with <i>paroxysmal</i> AF. We believe this may not be HERC's intention as such language would be counter to the prevailing clinical guidelines and clinical trial evidence cited in the coverage guidance document. The 2014 AHA/ACC/HRS Atrial Fibrillation Guideline (referenced on page 13 of the coverage guidance document) strongly recommends PVI for patients with symptomatic paroxysmal or symptomatic persistent AF who are refractory or intolerant to at least 1 class I or III antiarrhythmic medication (January 2014). We propose the following revised coverage language: <i>Transcatheter pulmonary vein isolation is recommended for coverage for those who are symptomatic (paroxysmal AF or persistent AF) despite rate control medications or antiarrhythmic medications (strong recommendation).</i>	EbGS agrees that the recommendation as written can be confusing and has modified the language to clarify: Transcatheter pulmonary vein isolation is recommended for coverage for those who ~~are persistently~~ remain symptomatic from atrial fibrillation despite rate control medications and antiarrhythmic medications (strong recommendation)
C	3	This revised language will help to ensure alignment with current clinical practice guidelines and that the appropriate patients have access to the right treatment strategies.	EbGS agrees.

HERC Coverage Guidance – Ablation for Atrial Fibrillation Disposition of Public Comments

References Provided by Commenters

Commenter	References
C	(1) January, C. T., Calkins, H., Murray, K. T., Cigarroa, J. E., & Stevenson, W. G. (2014). 2014 AHA/ACC/HRS Guideline for the Management of Patients with Atrial Fibrillation. <i>Circulation</i> , 129, 000-00.

COVERAGE GUIDANCE: NUCLEAR CARDIAC IMAGING

Question: How should the EBGs Coverage Guidance regarding nuclear cardiac imaging be applied to the Prioritized List?

Question source: Evidence Based Guideline Subcommittee

Issue: EBGs approved a new Coverage Guidance at their September, 2014 meeting. This CG needs final approval by HERC at the November, 2014 meeting. The summary of the coverage guidance is shown below.

HERC COVERAGE GUIDANCE

PET is not recommended for coverage for screening or diagnosis of coronary artery disease (CAD) (*strong recommendation*).

Single photon emission computed tomography (SPECT) is not recommended for coverage for screening for CAD in asymptomatic patients (*strong recommendation*).

SPECT is not recommended for coverage for diagnosis or risk stratification of CAD (*strong recommendation*)—except in patients for whom stress imaging is required and stress ECHO is contraindicated, is unavailable or would provide suboptimal imaging*)

**i.e. pre-existing cardiomyopathy, baseline regional wall motion abnormalities, left bundle branch block, paced rhythm, unsuitable acoustic windows due to body habitus, inability to utilize dobutamine in a setting where exercise is not possible or when the target workload is not achievable*

Evidence Summary

In **asymptomatic individuals** at high risk of CAD, there is no evidence of benefit for SPECT screening compared to no screening. In **symptomatic patients at low to intermediate risk of CAD**, evidence is conflicting with regard to ability to predict mortality and cardiovascular events, with one study finding no difference between ETT and SPECT, and another finding that stress SPECT and stress ECHO were better predictors than ETT and rest ECHO. In **symptomatic patients at high risk of CAD**, evidence is conflicting regarding rates of revascularization in those who undergo ETT compared to SPECT. Prognostic value does not differ between stress ECHO and stress SPECT. In **populations with mixed risk of CAD**, stress SPECT, stress ECHO, stress CMR and angiography do not differ in subsequent death or patient reported adverse cardiac events. SPECT and ECHO have similar prognostic abilities, and those tests as well as cardiac MR result in similar proportions of referrals to angiography or change in medical management.

With regard to diagnostic accuracy, SPECT and ECHO have similar sensitivity (83% to 87%) and specificity (64% to 77%), although some analyses suggest that ECHO may be slightly more sensitive and SPECT may be slightly more specific. Extracardiac findings (which may require additional evaluation) are identified rarely with SPECT, and significantly less frequently than CCTA.

Comparative evidence on the risks of various testing strategies is very limited, with the only apparent difference being that exercise stress has lower rates of adverse events than pharmacologic stress. SPECT has the highest radiation exposure of any testing strategy at a range of 7 to 30 mSv. SPECT appears to perform similarly in men and women, Caucasians and African-Americans, normal weight and obese patients, those with and without diabetes and those with and without hypertension.

Evidence is conflicting regarding the value of ETT compared to SPECT. One study suggests that direct referral to angiography is the most cost effective strategy, with SPECT and cardiac MR being of moderate cost, and ECHO being the most costly. Another analysis finds that SPECT has low value compared to no screening in an asymptomatic population and compared to ETT in a low-to-intermediate risk population, and has comparable value compared to ECHO in all other populations.

The evidence pertaining to PET is insufficient to draw conclusions for any outcome.

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variation in values and preferences	Coverage recommendation	Rationale
SPECT in asymptomatic at high risk of CAD	No net benefit, potential harm of radiation	High	Moderate	Low variability	Not recommended for coverage (strong recommendation)	Sufficient evidence of no net benefit, potential harms due to radiation, and higher cost.
SPECT in symptomatic with low/mod risk of CAD	Comparable to ETT and ECHO; potential harms from radiation	Moderate (ETT) High (ECHO)	Moderate	Moderate variability	Not recommended for coverage (strong recommendation)	Sufficient evidence of no net benefit, potential harms due to radiation. **
SPECT in symptomatic with high risk of CAD	Small health benefit compared to ETT, comparable to ECHO but potential harms from radiation	Moderate (ETT) High (ECHO)	Moderate	Moderate variability	<i>SPECT vs. ETT</i> Not recommended for coverage (weak recommendation) <i>SPECT vs. ECHO</i> Not recommended for coverage (strong recommendation)	<i>SPECT vs. ETT</i> Sufficient evidence of small net benefit over ETT, may outweigh potential harms due to radiation, upgrading algorithmic derived recommendation against from “strong” to “weak”. <i>SPECT vs. ECHO**</i> Sufficient evidence of no net benefit, potential harms due to radiation.
SPECT in known CAD	Unknown compared to ETT Comparable to ECHO but potential harms from radiation	Insufficient (ETT) High (ECHO)	Moderate	Moderate variability	Not recommended for coverage (strong recommendation)	<i>SPECT vs. ETT</i> Insufficient evidence, but has potential risks of radiation exposure, unlike alternatives, and is higher cost.

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variation in values and preferences	Coverage recommendation	Rationale
						<i>SPECT vs. ECHO</i> ** Sufficient evidence of no net benefit, potential harms due to radiation.
PET in all populations	Unknown compared to all interventions	Insufficient	High	Low variability	Not recommended for coverage (strong recommendation)	Insufficient evidence; also has potential risks of radiation exposure, unlike alternatives and is higher cost.

*The Quality of Evidence rating was assigned by the primary evidence source, not the HERC Subcommittee

**Expert input led to the decision to recommend coverage for SPECT when stress echo may be contraindicated or provide suboptimal imaging, in the following circumstances: pre-existing cardiomyopathy, baseline regional wall motion abnormalities, left bundle branch block, paced rhythm, unsuitable acoustic windows due to body habitus, inability to utilize dobutamine in a setting where exercise is not possible or when the target workload is not achievable

Current Prioritized List status

CODES	DESCRIPTION	Current Lines
402.0-402.9	Hypertensive heart disease	80 HYPERTENSION AND HYPERTENSIVE DISEASE
411.0	Postmyocardial infarction syndrome	73 ACUTE AND SUBACUTE ISCHEMIC HEART DISEASE, MYOCARDIAL INFARCTION
411.1	Intermediate coronary syndrome (impending infarction, preinfarction angina, preinfarction syndrome, unstable angina)	73
411.8	Other acute and subacute forms of ischemic heart disease	73
411.81	Acute coronary occlusion without myocardial infarction	73
411.89	Other acute and subacute forms of ischemic heart disease, other	73
413.0	Angina decubitus (nocturnal angina)	193 CHRONIC ISCHEMIC HEART DISEASE
413.1	Prinzmetal angina (variant angina pectoris)	193
413.9	Other and unspecified angina pectoris (NOS, cardiac, equivalent, of effort, angina syndrome, status anginosus, stenocardia, syncope anginosa)	193
428.0	Congestive heart failure, unspecified	102 HEART FAILURE 267 CONGESTIVE HEART FAILURE, CARDIOMYOPATHY, MALIGNANT ARRHYTHMIAS, AND COMPLEX CONGENITAL HEART DISEASE Treatment: CARDIAC TRANSPLANT
428.1	Left heart failure	102,267
428.20	Systolic heart failure, unspecified	102,267 286 LIFE-THREATENING CARDIAC ARRHYTHMIAS
428.21	Acute systolic heart failure	102,267,286
428.22	Chronic systolic heart failure	102,267,286
428.23	Acute on chronic systolic heart failure	102,267,286
428.30	Diastolic heart failure, unspecified	102,267,286
428.31	Acute diastolic heart failure	102,267,286
428.32	Chronic diastolic heart failure	102,267,286
428.33	Acute on chronic diastolic heart failure	102,267,286

CODES	DESCRIPTION	Current Lines
428.40	Combined systolic and diastolic heart failure, unspecified	102,267,286
428.41	Acute combined systolic and diastolic heart failure	102,267,286
428.42	Chronic combined systolic and diastolic heart failure	102,267,286
428.43	Acute on chronic combined systolic and diastolic heart failure	102,267,286
428.9	Heart failure, unspecified (cardiac failure, NOS, heart failure NOS, myocardial failure NOS, weak heart	102,267,286
429.2	Cardiovascular disease, unspecified	73,193
429.3	Cardiomegaly	662 CARDIOVASCULAR CONDITIONS WITH NO OR MINIMALLY EFFECTIVE TREATMENTS OR NO TREATMENT NECESSARY
429.4	Functional disturbances following cardiac surgery	286, 350 CARDIAC ARRHYTHMIAS
429.7	Certain sequelae of myocardial infarction not elsewhere classified	193
429.79	Certain sequelae of myocardial infarction not elsewhere classified, other	193
429.83	Takotsubo syndrome	102
429.9	Heart disease, unspecified	662
786.50	Chest pain, unspecified	Diagnostic Workup File
CPT Codes		
75557	Cardiac magnetic resonance imaging for morphology and function without contrast material	Congenital heart disease lines (48, 71,74,89,90,93,94,109, 110, 115, 132,134,142,180,192,236) 190 RHEUMATIC MULTIPLE VALVULAR DISEASE 227 DISEASES AND DISORDERS OF AORTIC VALVE 261 DISEASES OF MITRAL, TRICUSPID, AND PULMONARY VALVES 267 CONGESTIVE HEART FAILURE, CARDIOMYOPATHY, MALIGNANT ARRHYTHMIAS, AND COMPLEX CONGENITAL HEART DISEASE 662

CODES	DESCRIPTION	Current Lines
75559	with stress imaging	See 75557 above (but not 662)
75561	Cardiac magnetic resonance imaging for morphology and function without contrast material(s), followed by contrast material(s) and further sequences	See 75557 above
75563	with stress imaging	See 75557 above (but not 48)
75565	Cardiac magnetic resonance imaging for velocity flow mapping (List separately in addition to code for primary procedure)	See 75557 above
75571	Computed tomography, heart, without contrast material, with quantitative evaluation of coronary calcium	Excluded
75572	Computed tomography, heart, with contrast material, for evaluation of cardiac structure and morphology (including 3D image postprocessing, assessment of cardiac function, and evaluation of venous structures, if performed)	Excluded
75573	Computed tomography, heart, with contrast material, for evaluation of cardiac structure and morphology in the setting of congenital heart disease (including 3D image postprocessing, assessment of LV cardiac function, RV structure and function and evaluation of venous structures, if performed)	See 75557 above (but not 48)
78451	Myocardial perfusion imaging, tomographic (SPECT) (including attenuation correction, qualitative or quantitative wall motion, ejection fraction by first pass or gated technique, additional quantification, when performed); single study, at rest or stress (exercise or pharmacologic)	Diagnostic Procedure File
78452	multiple studies, at rest and/or stress (exercise or pharmacologic) and/or redistribution and/or rest reinjection	Diagnostic Procedure File
78453	Myocardial perfusion imaging, planar (including qualitative or quantitative wall motion, ejection fraction by first pass or gated technique, additional quantification, when performed); single study, at rest or stress (exercise or pharmacologic)	Diagnostic Procedure File
78454	multiple studies, at rest and/or stress (exercise or pharmacologic) and/or	Diagnostic Procedure File

CODES	DESCRIPTION	Current Lines
	redistribution and/or rest reinjection	
78459	Myocardial imaging, positron emission tomography (PET), metabolic evaluation	Excluded
78472	Cardiac blood pool imaging, gated equilibrium; planar, single study at rest or stress (exercise and/or pharmacologic), wall motion study plus ejection fraction, with or without additional quantitative processing	Diagnostic Procedure File
78473	multiple studies, wall motion study plus ejection fraction, at rest and stress (exercise and/or pharmacologic), with or without additional quantification	Diagnostic Procedure File
78481/3	Cardiac blood pool imaging (planar), first pass technique; single study, at rest or with stress (exercise and/or pharmacologic), wall motion study plus ejection fraction, with or without quantification (single/multiple)	Diagnostic Procedure File
78491	Myocardial imaging, positron emission tomography (PET), perfusion; single study at rest or stress	Excluded
78492	multiple studies at rest and/or stress	Excluded
78499	Unlisted cardiovascular procedure, diagnostic nuclear medicine	Diagnostic Procedure File

HERC staff recommendations:

- 1) Make no change in current non-coverage of cardiac PET scan
 - a. CPT 78459, 78491, and 78492 are Excluded
 - b. Will add entries to the non-covered table for these CPT codes
- 2) Keep SPECT on the Diagnostic List; adopt the new diagnostic guideline below for SPECT imaging

DIAGNOSTIC GUIDELINE DXX, SPECT

SPECT (CPT 78451, 78452) is not covered for screening for coronary artery disease in asymptomatic patients.

SPECT is only covered for diagnosis or risk stratification of coronary artery disease in patients for whom stress imaging is required and stress ECHO is contraindicated, is unavailable or would provide suboptimal imaging (i.e. pre-existing cardiomyopathy, baseline regional wall motion abnormalities, left bundle branch block, paced rhythm, unsuitable acoustic windows due to body habitus, inability to utilize dobutamine in a setting where exercise is not possible or when the target workload is not achievable).

HEALTH EVIDENCE REVIEW COMMISSION (HERC)
COVERAGE GUIDANCE: NUCLEAR CARDIAC IMAGING
DRAFT for VbBS meeting materials 11/13/2014

HERC COVERAGE GUIDANCE

PET is not recommended for coverage for screening or diagnosis of coronary artery disease (CAD) (*strong recommendation*).

Single photon emission computed tomography (SPECT) is not recommended for coverage for screening for CAD in asymptomatic patients (*strong recommendation*).

SPECT is not recommended for coverage for diagnosis or risk stratification of CAD (*strong recommendation*)—except in patients for whom stress imaging is required and stress ECHO is contraindicated, is unavailable or would provide suboptimal imaging*)

**i.e. pre-existing cardiomyopathy, baseline regional wall motion abnormalities, left bundle branch block, paced rhythm, unsuitable acoustic windows due to body habitus, inability to utilize dobutamine in a setting where exercise is not possible or when the target workload is not achievable*

Note: Definitions for strength of recommendation are provided in Appendix A GRADE Element Description

RATIONALE FOR GUIDANCE DEVELOPMENT

The HERC selects topics for guideline development or technology assessment based on the following principles:

- Represents a significant burden of disease
- Represents important uncertainty with regard to efficacy or harms
- Represents important variation or controversy in clinical care
- Represents high costs, significant economic impact
- Topic is of high public interest

Coverage guidance development follows to translate the evidence review to a policy decision. Coverage guidance may be based on an evidence-based guideline developed by the Evidence-based Guideline Subcommittee or a health technology assessment developed by the Health Technology Assessment Subcommittee. In addition, coverage guidance may utilize an existing evidence report produced by one of HERC's trusted sources, generally within the last three years.

EVIDENCE SOURCE

Washington State Health Care Authority Health Technology Assessment Program. (2013). *Cardiac Nuclear Imaging*. Olympia, WA: Health Technology Assessment Program. Retrieved December 2, 2013, from <http://www.hca.wa.gov/hta/Pages/nuclear.aspx>

The summary of evidence in this document is derived directly from this evidence source, and portions are extracted verbatim.

SUMMARY OF EVIDENCE

Clinical Background

Coronary artery disease (CAD) is among the most common chronic conditions in the U.S., affecting over 16 million adults. Due to its prevalence, and because several options (e.g., surgery, medication) exist to reduce CAD-related morbidity and mortality, accurate diagnosis and/or risk stratification of CAD is critical. Currently the definitive standard for diagnosis is invasive coronary angiography. Because angiography primarily documents the anatomic presence of significant stenosis rather than identifying the “culprit” lesions likely to cause an adverse cardiovascular event, a growing number of non-invasive tests have been developed to identify CAD lesions significant enough to affect the flow of blood to the heart (i.e., myocardial perfusion). These functional tests are typically performed under exercise- or pharmacologically induced stress to determine whether blood flow deteriorates when the stressor is introduced.

The most common tests of cardiac function include the stress-electrocardiogram (ECG), or treadmill test (ETT), which measures cardiac activity via electrical signals, and the echocardiogram (ECHO), which uses ultrasound to measure abnormalities in heart wall motion using 2-dimensional imagery. ETT has fallen out of favor for use in patients at higher risk of CAD, however, as it has relatively low sensitivity in these patients, while stress-ECHO has been found to lack precision in detecting single-vessel versus multi-vessel disease and may produce suboptimal imagery in obese patients, those with chronic respiratory conditions, and patients with chest deformities or pre-existing myocardial damage.

To address some of these concerns, “nuclear imaging tests” have been developed to provide perfusion data in a broader spectrum of patients. The most longstanding of these is single photon emission computed tomography (SPECT), which uses a radioactive tracer and gamma camera to obtain 3-dimensional images of tracer uptake; areas of poor uptake are associated with abnormal levels of perfusion. Positron emission tomography (PET) scanners are also used with a radiotracer, and are felt by some to provide better image resolution in heavier patients and those with dense breast

tissue. So-called “hybrid” modalities have also been introduced to visualize both perfusion abnormalities and anatomic lesions using CT or MRI imagery in addition to nuclear testing.

There are trends in the use of cardiac nuclear imaging tests that are currently points of controversy, however. For one, the use of nuclear imaging for cardiovascular testing has grown substantially in recent years. In addition, questions have been raised about the appropriateness of nuclear imaging in certain populations. A substantial decrease in the prevalence of abnormal findings on such tests has been observed over time, due in part to greater recognition and treatment of cardiac risk factors but also to possible changes in referral patterns. This combination of substantial growth in utilization of cardiac nuclear imaging and declining rates of “positive” test results raises questions about the populations and indications for which such testing is appropriate.

Evidence Review

In the Washington HTA report, “symptomatic” means a patient with symptoms suggestive of myocardial ischemia (symptoms not specified). Risk categories of low, moderate and high were defined by the authors based on the Diamond-Forrester model of pretest probability, which incorporates age, gender and type of chest pain. These categories equate to probability ranges of <10%, 10-90% and >90% respectively. However, when other risk classification systems were used in the included studies, that information was utilized and reported by the authors.

Comparative Clinical Effectiveness

Asymptomatic Patients at High Risk of CAD

The one available study assessing the impact of cardiac nuclear imaging in asymptomatic, high-risk patients found no difference between SPECT screening and no screening in mortality or cardiovascular events, although many patients in both groups received subsequent stress testing for clinical reasons over approximately 5 years of follow-up. SPECT screening did increase the short-term rates of referral for angiography and revascularization vs. no screening.

Symptomatic Patients at Low-to-Intermediate Risk of CAD

Correlation of Imaging Study Findings with Mortality and Cardiovascular Events

Rates of mortality and major cardiac adverse events (MACE) did not generally differ between imaging modalities in available studies. Patients in the WOMEN study, an RCT of 772 women randomized to SPECT or ETT-based testing strategies were at very low CAD risk. The rates of all major adverse cardiovascular events at 2 years were 1.7% and 2.3% for ETT and rest/stress SPECT respectively, but this difference was not

significant (Hazard Ratio [HR]: 1.3; 95% CI: 0.5, 3.5; $p=.59$). The rate of revascularization also did not statistically differ between groups.

The long-term prognostic value of exercise SPECT, exercise ECHO, ETT, and clinical parameters was measured in a single cohort of 248 patients who were followed for a mean of 3.7 years. A total of 64 MACE occurred during follow-up. In multivariate analyses examining the incremental impact of (1) clinical + ETT data; (2) data in (1) + rest ECHO data; (3) data in (1) + exercise ECHO data; and (4) data in (1) + exercise SPECT data on predicting MACE events, the area under the curve¹ did not statistically differ between the SPECT and ECHO models (0.78 and 0.77 respectively), but was significantly ($p<.05$) higher than the base model (0.68) or the rest ECHO model (0.72).

One study evaluated the impact on all-cause mortality of normal findings on stress-only vs. stress/rest SPECT protocols in nearly 17,000 low-to-intermediate risk patients followed for a median of 4.5 years. Annualized unadjusted mortality rates were statistically-significantly greater in the stress/rest group (2.92% vs. 2.57% for stress-only, $p=.02$); however, this difference was no longer apparent after multivariate adjustment for differences in baseline characteristics. The authors conclude that a stress/rest protocol may be unnecessary in lower-risk individuals.

Downstream Testing and Clinical Decision-Making

The impact of testing on downstream resource utilization and clinical decisions was evaluated only in the WOMEN study. Over 2 years of follow-up, repeat testing with the same modality was more frequent in the SPECT group vs. ETT (9% vs. 3%), although this difference was not statistically tested. However, 18% of women randomized to ETT crossed over to SPECT during follow-up. The overall rate of referral to angiography was higher in the ETT group (9.0% vs. 5.5% for SPECT, $p<.0001$). Changes in the use of nitrates, beta-blockers, and antidepressant therapies during follow-up did not differ between the two arms in the study.

Health-related Quality of Life

The impact of testing on health-related quality of life (HrQOL) also was examined only in the WOMEN study. Similar proportions of women in each treatment group reported “excellent” or “very good” QoL as well as “best” or “average” life satisfaction, with no statistical differences between groups. There were also no statistically-significant differences between ETT and SPECT groups in relation to changes in any of the subscales.

¹ This measure refers to the receiver operating characteristic (ROC) curve, a figure depicting the power of a diagnostic test. It includes both test sensitivity and specificity. A ROC curve for a perfect test has an area under the ROC curve of 1.0, while a test that performs no better than chance has an area under the curve of 0.5.

Symptomatic Patients at High Risk of CAD

Correlation of Imaging Study Findings with Mortality and Cardiovascular Events

In high risk populations, some differences in event rates by modality were apparent. An RCT of ETT vs. SPECT in 457 intermediate-to-high risk patients focused primarily on the period between testing and diagnosis, but did report on the rate of revascularization, which occurred more frequently in the ETT group (18% vs. 11% for SPECT, not statistically tested). In the “SPARC” registry, a study comparing short-term outcomes of PET, SPECT and coronary CT angiography (CCTA), revascularization rates at 90 days did not materially differ between PET and SPECT, regardless of whether findings were mildly or moderately-severely abnormal. Neither of these studies evaluated longer term outcomes such as mortality or cardiovascular events.

Another study assessed the prognostic value of both dobutamine ECHO and dobutamine SPECT in 301 patients who were unable to exercise and were at intermediate-to-high risk of CAD; patients were followed for a mean of 7.3 years. Event-free survival was significantly better for patients with normal vs. abnormal findings on both tests, and did not differ statistically between tests. In multivariate models based on clinical data, stress testing, and imaging results, abnormal findings on either SPECT or ECHO were the strongest predictors of both cardiac death (HR [95% CI]: 4.4 [1.2, 21.0] and 3.4 [1.2, 12.0] for SPECT and ECHO respectively) and cardiac events (3.1 [1.1, 8.9] and 2.6 [1.1-6.2] respectively).

Downstream Testing and Clinical Decision-Making

Two studies reported on the effects of testing on downstream resource use and/or clinical decisions. Of the 207 patients randomized to ETT in one RCT, a total of 146 (71%) were referred for further testing (47% to angiography and 23% to stress ECHO). In contrast, further testing was requested in only 16% of patients randomized to SPECT, all of which were angiography procedures ($p < .0001$ for the comparison). ETT also appeared to generate more false-positives for significant CAD. Only 38% of ETT patients referred to angiography were revascularized, vs. 66% of SPECT patients so referred ($p < .05$).

In a registry study, referral for angiography occurred in a greater percentage of PET patients (11.1% vs. 4.3% for SPECT; $p < .001$). In multivariate analyses controlling for patient characteristics, comorbidities, and testing location, imaging modality was significantly and positively correlated with referral to angiography for PET (OR: 5.0; 95% CI: 1.0, 24.4) in comparison to SPECT. Neither PET nor SPECT were associated with significant medication changes.

Health-Related Quality of Life

There were no studies in symptomatic, high-risk individuals that reported on the impact of cardiac nuclear imaging tests on HrQoL.

Known CAD

Correlation of Imaging Study Results with Mortality and Cardiovascular Events

One comparative cohort study compared the rate of revascularization in 2,951 patients with known CAD and left ventricular dysfunction and (1) who had been tested with SPECT before referral for angiography; (2) were tested with SPECT only after a positive angiography; or (3) had no SPECT before or after angiography. The rate of revascularization differed significantly ($p=.001$) among groups, with the lowest rate of 35.8% seen in postangiography SPECT patients, 45.6% in patients who had SPECT pre-angiography, and 53.2% among patients undergoing angiography with SPECT neither before nor afterward.

Downstream Testing and Clinical Decision-Making

In one study, 100 consecutive patients, 79% of whom had known CAD, underwent rest-stress PET perfusion testing. Physicians were first queried on proposed patient management strategies without PET perfusion data; actual patient management was measured 4 weeks after PET. Proposed patient management was altered in 78% of patients. Most prominently, conservative medical management was initially proposed in 28% of patients; after PET testing, 76% were managed this way in actuality. In addition, use of angiography to guide treatment via PTCA was proposed in 6%, but was performed in 20% after PET testing.

Health-Related Quality of Life

There were no studies in patients with known CAD that reported on the impact of cardiac nuclear imaging tests on HrQoL.

No comparative studies evaluating the impact of serial nuclear imaging in asymptomatic patients with known CAD were identified.

Mixed Populations

The largest number of studies was available for populations that did not fit neatly into the categories described above. They represented a true “mix” of patients based on relatively uniform distributions by risk or pretest probability, presence or absence of symptoms, and/or inclusion of patients with known vs. suspected CAD. A total of 10 studies were identified.

Correlation of Imaging Study Results with Mortality and Cardiovascular Events

Data on mortality and cardiovascular events were available in 8 studies. The Cost-Effectiveness of Functional Cardiac Testing (CeCAT) Trial was an RCT comparing multiple diagnostic strategies—rest-adenosine stress SPECT, ECHO (dobutamine stress), adenosine stress cardiac magnetic resonance imaging (cardiac MR), and direct referral to angiography—among 898 primarily high-risk patients with known or suspected CAD and stable symptoms of ischemia who were referred to a tertiary center

in the UK for angiography and were followed for 18 months. In this study, the number of total, cardiac, and noncardiac deaths did not statistically differ by imaging modality. When compared with the referent angiography group, the number of nonfatal adverse cardiac events did not differ for SPECT or cardiac MR, but was statistically-significantly higher for ECHO (relative risk [RR]: 1.95; 95% CI: 1.23, 3.08; $p=.012$), primarily because of more admissions for chest pain. When the number of *patients* reporting adverse cardiac events was compared, however, no significant differences were observed (one patient in the ECHO group was responsible for seven hospital admissions).

Findings from a study comparing PET and SPECT were somewhat mixed. No differences in cardiovascular mortality or the rate of MI were observed between groups. However, the rates of CABG (3.4% vs. 7.8%, $p<.01$) and any revascularization (6.0% vs. 11.4%, $p<.01$) were statistically-significantly lower for PET vs. the internal (identified by report authors) SPECT control group. The rate of any revascularization was also significantly lower in comparison to the external (using results from another published trial) SPECT control group (6.0% vs. 13.0%, $p<.0001$).

Three cohort studies comparing the prognostic ability of SPECT and ECHO generally showed comparable results for both tests. No statistical differences between imaging modalities in event rates or event-free survival were observed in 2 studies. In the third, an evaluation of exercise stress ECHO vs. exercise stress SPECT in 206 symptomatic veterans who received both tests and were followed for up to 10 years, moderate-to-large ischemia on ECHO was the strongest independent predictor of overall mortality (RR: 6.2; $p<.0001$), cardiovascular death (RR: 17.6; $p=.01$), congestive heart failure (RR: 17.4; $p=.0005$), or sudden death (RR: 26.8; $p=.003$). The presence of moderate-to-large fixed defects on SPECT was the strongest independent predictor of nonfatal MI (RR: 8.1; $p=.0002$) and unstable angina (RR: 3.0; $p=.005$).

One study assessed the predictive capability of functional data from ETT, exercise stress SPECT, and the "Gensini score" from angiography evaluation in 732 patients who were followed for a mean of 3.5 years. Abnormal results on SPECT and the Gensini score were significantly ($p\leq.01$) associated with poorer event-free survival, while ETT data were not. Analyses of the receiver operator curve (ROC) for events indicated that SPECT was the strongest independent predictor of events (0.67 vs. 0.61 and 0.46 for Gensini score and ETT, $p<.05$).

Downstream Testing and Clinical Decision-Making

A total of three studies examined the impact of cardiac nuclear imaging on further testing and clinical decision-making. In the CeCAT trial, the proportions of patients in the SPECT, ECHO, and cardiac MR groups who were referred to angiography ranged between 75-80% and did not statistically differ between groups; in addition, decisions on

further invasive or medical management were also similar. The rate of referral to angiography in the study comparing PET and SPECT was statistically-significantly lower for PET (13%) in comparison to both the internal (identified by report authors) and external (using results from another published trial) SPECT groups (31% and 34% respectively, $p<.0001$). The rate of angiography negative results was also significantly lower for PET vs. internal SPECT controls (5.2% vs. 15.6%, $p<.0001$).

Finally, a hypothetical referral rate to angiography was assessed in 955 patients undergoing ETT and rest-exercise stress SPECT. Algorithms using ETT data alone, SPECT data alone, and a combination of the 2 tests were applied. An estimated 27% of patients would have been referred to angiography based on ETT results alone, vs. 13% for SPECT data alone and 12% using both ETT and SPECT data ($p<.01$ for both comparisons to ETT alone). Findings were similar when compared among patients without known CAD.

Health-Related Quality of Life

HrQoL was assessed in the CeCAT trial. While some statistically-significant differences were noted in certain subscales at particular time points, improvements in HrQoL were clinically comparable across testing groups for all measures.

Diagnostic Accuracy

A total of 8 studies were available that examined the accuracy of cardiac nuclear imaging tests in relation to a functional reference standard. This is currently believed to be a more accurate method to determine whether a defect noted on non-invasive imaging relates to CAD that is functionally-significant—that is, likely to be the cause of an adverse cardiovascular event if not treated. Recent research has raised questions about the use of anatomic data on angiography to confirm findings of functional tests such as ECHO, SPECT, and PET. There is nevertheless a large body of evidence evaluating the accuracy of noninvasive functional tests using visualization of coronary arteries as the reference standard.

One of the most widely-cited meta-analyses compared the diagnostic accuracy of exercise ECHO and exercise SPECT based on 44 studies. Pooled sensitivity of the 2 tests was similar (85% and 87% for ECHO and SPECT respectively), but pooled specificity was rated higher for ECHO (77% vs. 64% for SPECT, $p<.05$). However, substantial heterogeneity in study populations, imaging protocols, and SPECT radiotracers was noted for this sample; subsequent reanalysis with controls for heterogeneity found no statistical differences between the tests.

Methods to assess diagnostic accuracy have also evolved, and feature newer techniques designed to capture the natural correlation between sensitivity and specificity. A recent meta-analysis using newer bivariate methods found that ECHO was

slightly more sensitive than SPECT (87% vs. 83% respectively), while SPECT was somewhat more specific (77% vs. 72% for ECHO). An additional bivariate meta-analysis using a much larger set of 113 SPECT studies found greater sensitivity (88%) and similar specificity (76%), although other commentators have noted that the older SPECT studies included in this review were subject to “verification bias” (i.e., use of the reference standard only in test-positive or other selected individuals), which tends to inflate sensitivity and may also reduce specificity. This meta-analysis also included estimates of diagnostic accuracy from 9 PET studies (pooled estimates of 93% and 81% for sensitivity and specificity respectively).

Finally, a third recent meta-analysis estimated diagnostic performance from 114 SPECT and 15 PET studies. SPECT sensitivity was similar to that reported elsewhere (88%), but specificity was somewhat lower (61%). Sensitivity and specificity for PET was estimated to be 84% and 81% respectively.

Other Outcomes

Extracardiac Findings

With the enhanced imagery available for many noninvasive tests, incidental findings outside of the area of interest can be problematic given the additional resources required for investigation. The reported rate of incidental extracardiac findings is very low with nuclear imaging tests given the limited field of detection, however; most available studies are limited to case reports of mediastinal masses. One recent study compared the rate of such findings between CCTA and SPECT in 479 patients; extracardiac findings requiring further investigation were detected in 7% of CCTA patients but in no SPECT patients ($p=.0001$). Another analysis examined images of 2,155 patients undergoing SPECT studies, 6 (0.3%) of whom had extracardiac findings requiring follow-up. Four of the 6 patients had malignancies requiring further treatment. No PET studies reported on extracardiac findings.

While SPECT itself is associated with a low rate of extracardiac findings, the increasing use of CT for attenuation correction may result in increased detection of these findings. In a cohort study assessing prevalence of extracardiac findings from 582 SPECT/CT studies, a total of 400 (68.7%) included noncardiac findings, 196 (33.7%) of which were felt to be potentially relevant.

Equivocal/Indeterminate Results

While equivocal or indeterminate findings are possible with any diagnostic test, these results are rarely published. A recent systematic review of nearly 1,200 diagnostic accuracy studies found that 35% reported the presence of inconclusive results. Inconclusive results were reported in only one of the studies in this report. In the CeCAT

trial comparing SPECT with ECHO, cardiac MR, and angiography, rates of equivocal findings were 4.0%, 6.6%, 6.6% and 2.0% respectively.

Risks of Testing

Patients appear to be at minimal immediate risk from cardiac nuclear imaging tests in and of themselves, although harms data are reported in only a small number of comparative studies. The risks that are reported are related primarily to the stressor employed (i.e., exercise or pharmacologic stress).

Comparative Data on Testing Risks

Only 2 studies compared adverse effects of multiple testing modalities. In the WOMEN study that randomized patients to ETT or exercise SPECT, no statistically significant differences between groups were noted in rates of chest pain, dyspnea, or fatigue after testing. In the CeCAT trial comparing SPECT, ECHO, cardiac MR, and angiography, specific reasons for failed tests were recorded. Failure to complete the test due to adverse effects occurred in 4 ECHO patients (1.8%), due to vasovagal reactions, blood pressure changes and dyspnea; no patient failed to complete SPECT due to adverse effects.

Adverse Effects by Stressor

Information on adverse effects attributed to specific stressors was obtained from 15 studies. Regardless of the comparisons made, events were typically described as non-serious and resolved once the stressor infusion ended. Reported ranges of adverse effects were similar across pharmacologic agents. Limited data suggest lower rates of adverse effects for exercise vs. pharmacologic stress in the 2 studies making this comparison, although statistical comparisons were not available for all event types.

Radiation Exposure

Potential adverse health effects associated with radiation exposure are important factors to consider in the evaluation of cardiac nuclear imaging tests, particularly because patients may already be exposed to radiation at other points along the diagnostic pathway (e.g., CCTA, angiography), cumulative radiation dose may be substantial in patients receiving serial imaging studies, and imaging alternatives such as ECHO and cardiac MR exist that do not involve radiation. Radiation dose is a measure of ionizing energy absorbed per unit of mass, expressed as units of Gy (Gray) or mGy; it often is quoted as an equivalent “effective” dose to major organs in the scanned area, in units of Sv (Sievert) or mSv. For x-rays, the radiation type produced by CT scanners, 1 mSv = 1 mGy. Average total effective dosages for SPECT range from 7 to 30 mSv, while for PET and CCTA the range is 2 to 14 mSv, and for invasive coronary angiography the range for is 5 to 7. While exposure to ionizing radiation at these levels

is associated with potential increase in cancer risk, the latency period for the development of such cancers may range from 10 to 40 years for solid tumors depending on the age and sex of the patient being tested. The intended use of cardiac imaging tests then becomes a critical consideration.

Differential Effectiveness/Safety for Key Patient Subgroups

The *comparative* impact of cardiac nuclear imaging tests vs. alternative testing strategies in certain subgroups is presented below.

Patient Age, Sex, Race or Ethnicity, and Comorbidities

A single comparative cohort study was available that assessed all-cause mortality for stress only vs. stress-rest SPECT (n=16,854) in specific subgroups over a mean of 4.5 years of follow-up. On a univariate basis, stress-rest protocols were associated with a statistically-significantly higher mortality rate in older (age >65) individuals, men, patients with a BMI <30 kg/m², and patients with diabetes. However, after multivariate adjustment for baseline characteristics, no statistically-significant differences remained. Several large cohort studies and meta-analyses have assessed the performance of SPECT in certain patient subgroups. For example, several studies have found that SPECT's diagnostic and prognostic performance is similar for women and men. Comparable results have also been found in several large ECHO studies. A meta-analysis of risk-stratification studies in over 13,000 patients age >65 years found that both stress SPECT and stress ECHO accurately risk-stratified patients vs. ETT. A multicenter cohort study of approximately 1,100 patients found that SPECT results were predictive of cardiac events in both Caucasian and African-American patients.

Analyses comparing patients with and without diabetes suggest that, while diabetes is a predictor of mortality for any nuclear imaging result, SPECT testing provides incremental prognostic information in patients with and without diabetes alike. Multiple studies have found that SPECT is feasible and has comparable diagnostic and prognostic performance in normal-weight, overweight, and obese patients. Finally, a meta-analysis of SPECT and ECHO studies in hypertensive patients showed diagnostic accuracy similar to that observed in all patients with suspicion of CAD.

Clinical Setting

In a comparison of stress-only vs. stress-rest SPECT, mortality was initially statistically-significantly higher in stress-rest patients in an inpatient setting. After multivariate adjustment, however, no significant differences remained. Limited additional data are available explicitly comparing the performance of SPECT by setting. One study evaluating the potential benefit of an emergency department chest pain clinic estimated that unnecessary hospitalizations would be reduced in 30% of patients and inappropriate discharges avoided in 6% through the use of a selective SPECT protocol.

Selection of Test by Primary Care vs. Specialty Physician

No study assessed the impact of ordering specialty on patient outcomes, clinical decision-making, or costs. There are, however, several studies that have assessed the impact of specialty on whether ordered cardiac SPECT studies meet published appropriate use criteria (AUC). In a multicenter assessment of an online SPECT appropriateness classification system, one study found that the rate of inappropriate studies was statistically-significantly higher among non-cardiologists (19.5% vs. 13.2% for cardiologists, $p < .0001$). Similar findings have been observed in several single center studies. Of note, most inappropriate ordering of SPECT perfusion studies appears to have occurred in women, younger patients, and/or those without symptoms.

Scan Vendor, Type of Assessment, Type of Radioisotope, and Type of Stressor

No study assessed the impact of scan vendor or qualitative vs. quantitative assessment on patient outcomes, clinical decision-making, or costs. Most of the studies evaluating differences according to stressor type focused on rates of adverse effects of pharmacologic testing. The study that evaluated stress-only vs. stress-rest SPECT found no statistically-significant effects on mortality with subgroups defined by exercise vs. pharmacologic stress.

Two studies examined the impact of different SPECT radiotracers on outcomes. In one, a total of 1,818 patients underwent exercise or pharmacologic stress SPECT with Tc-99m sestamibi or Tc-99m tetrofosmin. Patients were followed for a mean of 1.5 years, during which no statistically-significant differences were observed between groups in the rates of overall mortality, cardiovascular mortality, or the composite endpoint of cardiovascular mortality or nonfatal MI.

The other study compared mortality outcomes among 2,147 patients with known CAD undergoing pharmacologic stress SPECT with either Tc-99m sestamibi or Tc-99m tetrofosmin who were followed for a median of 4 years. During follow-up, a total of 704 all-cause deaths (493 cardiovascular-related) were reported. There was no significant difference in either overall or cardiovascular mortality between radiotracer groups on both an unadjusted and multivariate-adjusted basis.

Analysis of Comparative Value

Limited evidence is available that directly measured and compared the economic impact of non-invasive testing strategies for CAD. Three RCTs compared costs of SPECT to other imaging. In the only economic study performed in the US, an RCT of ETT vs. SPECT in 772 women at low-to-intermediate risk of CAD in 43 cardiology practices across the U.S., total mean costs of testing over 2 years were higher in the SPECT arm (\$643 vs. \$338, $p < .001$), as the higher costs of initial SPECT testing outweighed the increased costs of downstream testing in the ETT arm. In another 2-year RCT

conducted in 457 primarily intermediate-risk patients in the UK, however, downstream testing costs were substantially higher in the ETT arm, leading to significantly higher total costs from randomization to diagnosis using National Health Service (NHS) estimates (\$1,244 v \$743 for SPECT, $p < .001$). The final UK RCT compared costs of initial and repeat testing, treatment, and adverse events over 18 months of follow-up for mixed-risk patients randomized to SPECT, ECHO, cardiac MR, or direct referral to angiography. Direct referral to angiography was the lowest-cost strategy. Incremental costs (relative to angiography) were similar for the SPECT and cardiac MR strategies (~\$650), but were twice as high for patients in the ECHO group (~\$1,250) due to a higher rate of hospital readmissions.

Economic evidence for PET was limited to 2 studies. In one, an evaluation of planned vs. actual management before and after PET perfusion testing in 100 patients with known CAD, savings from reduced need for angiography were greater than the incremental costs of PET testing and revascularization, leading to overall savings of \$240 per patient. In the other, a matched comparative cohort analysis of PET and SPECT, mean costs of all diagnostic testing were approximately \$2,500 in both groups, but greater requirements for revascularization at 1 year led to higher total costs in the SPECT group (\$5,937 vs. \$4,110 for PET).

Because evidence is limited comparing the short-term clinical consequences and costs for all relevant non-invasive strategies for CAD diagnosis, the authors of this report developed a decision-analytic model to provide additional information. The target population involved men and women with suspected or known CAD who had stable symptoms of myocardial ischemia (i.e., atypical or typical chest pain or other symptoms such as dyspnea). Model outcomes and costs were estimated over a 90-day period. The authors of the Washington HTA report developed 7 different strategies, alone and in combination, to capture a wide range of management approaches:

1. ECHO, followed by invasive coronary angiography if ECHO is positive or inconclusive
2. ETT, followed by angiography if ETT is positive or inconclusive
3. SPECT, followed by angiography if ETT is positive or inconclusive
4. PET, followed by angiography if ETT is positive or inconclusive
5. ETT, followed by ECHO if ETT is positive or inconclusive, followed by angiography if the ECHO is positive or inconclusive
6. ETT, followed by SPECT if ETT is positive or inconclusive, followed by angiography if the SPECT is positive or inconclusive
7. ETT, followed by PET if ETT is positive or inconclusive, followed by angiography if the PET is positive or inconclusive

Because the underlying CAD prevalence varies in different patient populations, the authors calculate results of the identical testing strategies for a population with 10%, 30%, 50% and 70% CAD prevalence. Comparing these results demonstrates the importance of the underlying prevalence on the relative balance of false negatives, false positives, rates of referral to angiography, and costs. For example, among a patient population with a CAD prevalence of 10%, the difference in false negatives between SPECT and ECHO almost vanishes (4 per 1,000). In contrast, the difference in false positives between SPECT and ECHO in a population with 50% CAD prevalence was 33 per 1,000 but is increased to 60 per 1,000 when the underlying prevalence of CAD is only 10%. The relative differences in angiography referral, patients exposed to radiation, and costs also shift.

The authors of this report have devised their own evidence rating system, and reach the following conclusions for specific populations:

- Asymptomatic, high-risk individuals
 - o SPECT vs. no screening – high certainty of a comparable net health benefit, low value
 - o SPECT vs. ETT or ECHO – insufficient evidence
 - o PET vs. any alternative – insufficient evidence
- Symptomatic individuals at low-to-intermediate CAD risk
 - o SPECT vs. ETT – moderate certainty of a comparable net health benefit, low value
 - o SPECT vs. ECHO – high certainty of a comparable net health benefit, reasonable/comparable value
 - o PET vs. any alternative – insufficient evidence
- Symptomatic individuals at high CAD risk
 - o SPECT vs. ETT – moderate certainty of a small net health benefit, reasonable/comparable value
 - o SPECT vs. ECHO – high certainty of a comparable net health benefit, comparable/reasonable value
 - o PET vs. any alternative – insufficient evidence
- Known CAD
 - o SPECT vs. ETT – insufficient evidence
 - o SPECT vs. ECHO – high certainty of a comparable net health benefit, comparable/reasonable value

- o PET vs. any alternative – insufficient evidence

[\[Evidence Source\]](#)

Evidence Summary

In **asymptomatic individuals** at high risk of CAD, there is no evidence of benefit for SPECT screening compared to no screening. In **symptomatic patients at low to intermediate risk of CAD**, evidence is conflicting with regard to ability to predict mortality and cardiovascular events, with one study finding no difference between ETT and SPECT, and another finding that stress SPECT and stress ECHO were better predictors than ETT and rest ECHO. In **symptomatic patients at high risk of CAD**, evidence is conflicting regarding rates of revascularization in those who undergo ETT compared to SPECT. Prognostic value does not differ between stress ECHO and stress SPECT. In **populations with mixed risk of CAD**, stress SPECT, stress ECHO, stress CMR and angiography do not differ in subsequent death or patient reported adverse cardiac events. SPECT and ECHO have similar prognostic abilities, and those tests as well as cardiac MR result in similar proportions of referrals to angiography or change in medical management.

With regard to diagnostic accuracy, SPECT and ECHO have similar sensitivity (83% to 87%) and specificity (64% to 77%), although some analyses suggest that ECHO may be slightly more sensitive and SPECT may be slightly more specific. Extracardiac findings (which may require additional evaluation) are identified rarely with SPECT, and significantly less frequently than CCTA.

Comparative evidence on the risks of various testing strategies is very limited, with the only apparent difference being that exercise stress has lower rates of adverse events than pharmacologic stress. SPECT has the highest radiation exposure of any testing strategy at a range of 7 to 30 mSv. SPECT appears to perform similarly in men and women, Caucasians and African-Americans, normal weight and obese patients, those with and without diabetes and those with and without hypertension.

Evidence is conflicting regarding the value of ETT compared to SPECT. One study suggests that direct referral to angiography is the most cost effective strategy, with SPECT and cardiac MR being of moderate cost, and ECHO being the most costly. Another analysis finds that SPECT has low value compared to no screening in an asymptomatic population and compared to ETT in a low-to-intermediate risk population, and has comparable value compared to ECHO in all other populations.

The evidence pertaining to PET is insufficient to draw conclusions for any outcome.

GRADE-INFORMED FRAMEWORK

The HERC develops recommendations by using the concepts of the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system. GRADE is a transparent and structured process for developing and presenting evidence and for carrying out the steps involved in developing recommendations. There are four elements that determine the strength of a recommendation, as listed in the table below. The HERC reviews the evidence and makes an assessment of each element, which in turn is used to develop the recommendations presented in the coverage guidance box. Balance between desirable and undesirable effects, and quality of evidence, are derived from the evidence presented in this document, while estimated relative costs, values and preferences are assessments of the HERC members.

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variation in values and preferences	Coverage recommendation	Rationale
SPECT in asymptomatic at high risk of CAD	No net benefit, potential harm of radiation	High	Moderate	Low variability	Not recommended for coverage (strong recommendation)	Sufficient evidence of no net benefit, potential harms due to radiation, and higher cost.
SPECT in symptomatic with low/mod risk of CAD	Comparable to ETT and ECHO; potential harms from radiation	Moderate (ETT) High (ECHO)	Moderate	Moderate variability	Not recommended for coverage (strong recommendation)	Sufficient evidence of no net benefit, potential harms due to radiation. **
SPECT in symptomatic with high risk of CAD	Small health benefit compared to ETT, comparable to ECHO but potential harms from radiation	Moderate (ETT) High (ECHO)	Moderate	Moderate variability	<i>SPECT vs. ETT</i> Not recommended for coverage (weak recommendation) <i>SPECT vs. ECHO</i> Not recommended for coverage (strong recommendation)	<i>SPECT vs. ETT</i> Sufficient evidence of small net benefit over ETT, may outweigh potential harms due to radiation, upgrading algorithmic derived recommendation against from “strong” to “weak”. <i>SPECT vs. ECHO**</i> Sufficient evidence of no net benefit, potential harms due to radiation.

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variation in values and preferences	Coverage recommendation	Rationale
SPECT in known CAD	Unknown compared to ETT Comparable to ECHO but potential harms from radiation	Insufficient (ETT) High (ECHO)	Moderate	Moderate variability	Not recommended for coverage (strong recommendation)	<i>SPECT vs. ETT</i> Insufficient evidence, but has potential risks of radiation exposure, unlike alternatives, and is higher cost. <i>SPECT vs. ECHO**</i> Sufficient evidence of no net benefit, potential harms due to radiation.
PET in all populations	Unknown compared to all interventions	Insufficient	High	Low variability	Not recommended for coverage (strong recommendation)	Insufficient evidence; also has potential risks of radiation exposure, unlike alternatives and is higher cost.

*The Quality of Evidence rating was assigned by the primary evidence source, not the HERC Subcommittee

**Expert input led to the decision to recommend coverage for SPECT when stress echo may be contraindicated or provide suboptimal imaging, in the following circumstances: pre-existing cardiomyopathy, baseline regional wall motion abnormalities, left bundle branch block, paced rhythm, unsuitable acoustic windows due to body habitus, inability to utilize dobutamine in a setting where exercise is not possible or when the target workload is not achievable

Note: GRADE framework elements are described in Appendix A

POLICY LANDSCAPE

There were no quality measures that pertained to this topic identified when searching the [National Quality Measures Clearinghouse](#).

Coverage guidance is prepared by the Health Evidence Review Commission (HERC), HERC staff, and subcommittee members. The evidence summary is prepared by the Center for Evidence-based Policy at Oregon Health & Science University (the Center). This document is intended to guide public and private purchasers in Oregon in making informed decisions about health care services.

The Center is not engaged in rendering any clinical, legal, business or other professional advice. The statements in this document do not represent official policy positions of the Center. Researchers involved in preparing this document have no affiliations or financial involvement that conflict with material presented in this document.

Appendix A. GRADE Element Descriptions

Element	Description
Balance between desirable and undesirable effects	The larger the difference between the desirable and undesirable effects, the higher the likelihood that a strong recommendation is warranted. The narrower the gradient, the higher the likelihood that a weak recommendation is warranted
Quality of evidence	The higher the quality of evidence, the higher the likelihood that a strong recommendation is warranted
Resource allocation	The higher the costs of an intervention—that is, the greater the resources consumed—the lower the likelihood that a strong recommendation is warranted
Values and preferences	The more values and preferences vary, or the greater the uncertainty in values and preferences, the higher the likelihood that a weak recommendation is warranted

Strong recommendation

In Favor: The subcommittee is confident that the desirable effects of adherence to a recommendation outweigh the undesirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences.

Against: The subcommittee is confident that the undesirable effects of adherence to a recommendation outweigh the desirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences.

Weak recommendation

In Favor: The subcommittee concludes that the desirable effects of adherence to a recommendation probably outweigh the undesirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences, but is not confident.

Against: The subcommittee concludes that the undesirable effects of adherence to a recommendation probably outweigh the desirable effects, considering the quality of evidence, cost and resource allocation, and values and preferences, but is not confident.

Quality or strength of evidence rating across studies for the treatment/outcome²

High: The subcommittee is very confident that the true effect lies close to that of the estimate of the effect. Typical sets of studies are RCTs with few or no limitations and the estimate of effect is likely stable.

Moderate: The subcommittee is moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different. Typical sets of studies are RCTs with some limitations or well-performed nonrandomized studies with additional strengths that guard against potential bias and have large estimates of effects.

Low: The subcommittee's confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect. Typical sets of studies are RCTs with serious limitations or nonrandomized studies without special strengths.

Very low: The subcommittee has very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect. Typical sets of studies are nonrandomized studies with serious limitations or inconsistent results across studies.

² Includes risk of bias, precision, directness, consistency and publication bias

Appendix B. Applicable Codes

CODES	DESCRIPTION
ICD-9 Diagnosis Codes	
402.0-402.9	Hypertensive heart disease
411.0	Postmyocardial infarction syndrome
411.1	Intermediate coronary syndrome (impending infarction, preinfarction angina, preinfarction syndrome, unstable angina)
411.8	Other acute and subacute forms of ischemic heart disease
411.81	Acute coronary occlusion without myocardial infarction
411.89	Other acute and subacute forms of ischemic heart disease, other
413.0	Angina decubitus (nocturnal angina)
413.1	Prinzmetal angina (variant angina pectoris)
413.9	Other and unspecified angina pectoris (NOS, cardiac, equivalent, of effort, angina syndrome, status anginosus, stenocardia, syncope anginosa)
428.0	Congestive heart failure, unspecified
428.1	Left heart failure
428.20	Systolic heart failure, unspecified
428.21	Acute systolic heart failure
428.22	Chronic systolic heart failure
428.23	Acute on chronic systolic heart failure
428.30	Diastolic heart failure, unspecified
428.31	Acute diastolic heart failure
428.32	Chronic diastolic heart failure
428.33	Acute on chronic diastolic heart failure
428.40	Combined systolic and diastolic heart failure, unspecified
428.41	Acute combined systolic and diastolic heart failure
428.42	Chronic combined systolic and diastolic heart failure
428.43	Acute on chronic combined systolic and diastolic heart failure
428.9	Heart failure, unspecified (cardiac failure, NOS, heart failure NOS, myocardial failure NOS, weak heart)
429.2	Cardiovascular disease, unspecified
429.3	Cardiomegaly
429.4	Functional disturbances following cardiac surgery
429.7	Certain sequelae of myocardial infarction not elsewhere classified
429.79	Certain sequelae of myocardial infarction not elsewhere classified, other
429.83	Takotsubo syndrome
429.9	Heart disease, unspecified
786.50	Chest pain, unspecified
ICD-9 Volume 3 (Procedure Codes)	
92.05	Cardiovascular scan and radioisotope function study
CPT Codes	
75557	Cardiac magnetic resonance imaging for morphology and function without

CODES	DESCRIPTION
	contrast material
75559	with stress imaging
75561	Cardiac magnetic resonance imaging for morphology and function without contrast material(s), followed by contrast material(s) and further sequences
75563	with stress imaging
75565	Cardiac magnetic resonance imaging for velocity flow mapping (List separately in addition to code for primary procedure)
75571	Computed tomography, heart, without contrast material, with quantitative evaluation of coronary calcium
75572	Computed tomography, heart, with contrast material, for evaluation of cardiac structure and morphology (including 3D image postprocessing, assessment of cardiac function, and evaluation of venous structures, if performed)
75573	Computed tomography, heart, with contrast material, for evaluation of cardiac structure and morphology in the setting of congenital heart disease (including 3D image postprocessing, assessment of LV cardiac function, RV structure and function and evaluation of venous structures, if performed)
78451	Myocardial perfusion imaging, tomographic (SPECT) (including attenuation correction, qualitative or quantitative wall motion, ejection fraction by first pass or gated technique, additional quantification, when performed); single study, at rest or stress (exercise or pharmacologic)
78452	multiple studies, at rest and/or stress (exercise or pharmacologic) and/or redistribution and/or rest reinjection
78453	Myocardial perfusion imaging, planar (including qualitative or quantitative wall motion, ejection fraction by first pass or gated technique, additional quantification, when performed); single study, at rest or stress (exercise or pharmacologic)
78454	multiple studies, at rest and/or stress (exercise or pharmacologic) and/or redistribution and/or rest reinjection
78459	Myocardial imaging, positron emission tomography (PET), metabolic evaluation
78472	Cardiac blood pool imaging, gated equilibrium; planar, single study at rest or stress (exercise and/or pharmacologic), wall motion study plus ejection fraction, with or without additional quantitative processing
78473	multiple studies, wall motion study plus ejection fraction, at rest and stress (exercise and/or pharmacologic), with or without additional quantification
78481/3	Cardiac blood pool imaging (planar), first pass technique; single study, at rest or with stress (exercise and/or pharmacologic), wall motion study plus ejection fraction, with or without quantification (single/multiple)
78491	Myocardial imaging, positron emission tomography (PET), perfusion; single study at rest or stress
78492	multiple studies at rest and/or stress
78499	Unlisted cardiovascular procedure, diagnostic nuclear medicine

CODES	DESCRIPTION
93000	Electrocardiogram, routine ECG with at least 12 leads; with interpretation and report
93005	tracing only, without interpretation and report
93010	interpretation and report only
93015	Cardiovascular stress test using maximal or submaximal treadmill or bicycle exercise, continuous electrocardiographic monitoring, and/or pharmacological stress; with supervision, interpretation and report
93016	supervision only, without interpretation and report
93017	tracing only, without interpretation and report
93018	interpretation and report only
93350	Echocardiography, transthoracic, real-time with image documentation (2D), includes M-mode recording, when performed, during rest and cardiovascular stress test using treadmill, bicycle exercise and/or pharmacologically induced stress, with interpretation and report
93351	including performance of continuous electrocardiographic monitoring, with supervision by a physician or other qualified health care professional
93454	Catheter placement in coronary arter(s) for coronary angiography, including intraprocedural injection(s) for coronary angiography, imaging supervision and interpretation
93455	with catheter placement(s) in bypass graft(s) (internal mammary, free arterial, venous grafts) including intraprocedural injection(s) for bypass graft angiography
93456	with right heart catheterization
HCPCS Level II Codes	
A9500	Technetium tc-99m sestamibi, diagnostic, per study dose
A9502	Technetium tc-99m tetrofosmin, diagnostic, per study dose
A9505	Thallium tl-201 thallous chloride, diagnostic, per millicurie
A9526	Nitrogen n-13 ammonia, diagnostic, per study dose, up to 40 millicuries
A9555	Rubidium rb-82, diagnostic, per study dose, up to 60 millicuries
A9560	Technetium tc-99m labeled red blood cells, diagnostic, per study dose, up to 30 millicuries
J0280	Injection, aminophyllin, up to 250 mg
J0461	Injection, atropine sulfate, 0.01 mg
J0151	Injection, adenosine for diagnostic use, 1 mg
J1245	Injection, dipyridamole, per 10 mg
J1250	Injection, dobutamine hydrochloride, per 250 mg
J2785	Injection, regadenoson, 0.1 mg

Note: Inclusion on this list does not guarantee coverage

Appendix C. HERC Guidance Development Framework

HERC Guidance Development Framework Principles

This framework was developed to assist with the decision making process for the Oregon policy-making body, the HERC and its subcommittees. It is a general guide, and must be used in the context of clinical judgment. It is not possible to include all possible scenarios and factors that may influence a policy decision in a graphic format. While this framework provides a general structure, factors that may influence decisions that are not captured on the framework include but are not limited to the following:

- Estimate of the level of risk associated with the treatment, or any alternatives;
- Which alternatives the treatment should most appropriately be compared to;
- Whether there is a discrete and clear diagnosis;
- The definition of clinical significance for a particular treatment, and the expected margin of benefit compared to alternatives;
- The relative balance of benefit compared to harm;
- The degree of benefit compared to cost; e.g., if the benefit is small and the cost is large, the committee may make a decision different than the algorithm suggests;
- Specific indications and contraindications that may determine appropriateness;
- Expected values and preferences of patients.

SPECT in: Asymptomatic at high risk of CAD; Symptomatic with low/mod risk of CAD (compared to ETT and ECHO); Symptomatic with high risk of CAD (compared to ECHO); Known CAD (compared to ECHO)



Center for Evidence-based Policy

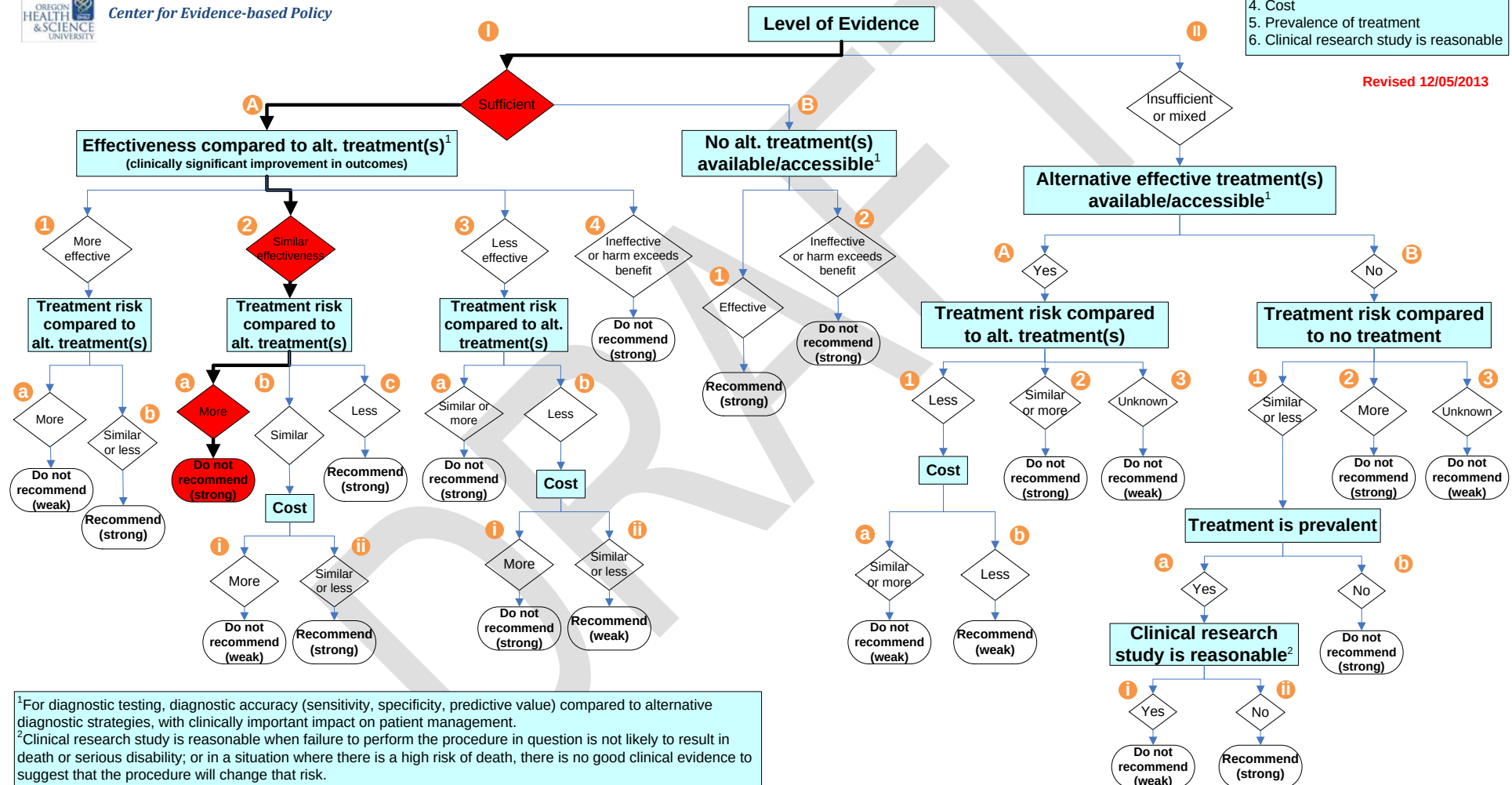
HERC Guidance Development Framework

Refer to *HERC Guidance Development Framework Principles* for additional considerations

Decision Point Priorities

1. Level of evidence
2. Effectiveness & alternative treatments
3. Harms and risk
4. Cost
5. Prevalence of treatment
6. Clinical research study is reasonable

Revised 12/05/2013

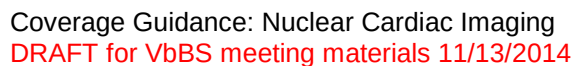




OREGON
HEALTH
& SCIENCE
UNIVERSITY

HERC Guidance Development Framework

Refer to *HERC Guidance Development Framework Principles* for additional considerations



SPECT in known CAD (compared to ETT); PET in all populations



Center for Evidence-based Policy

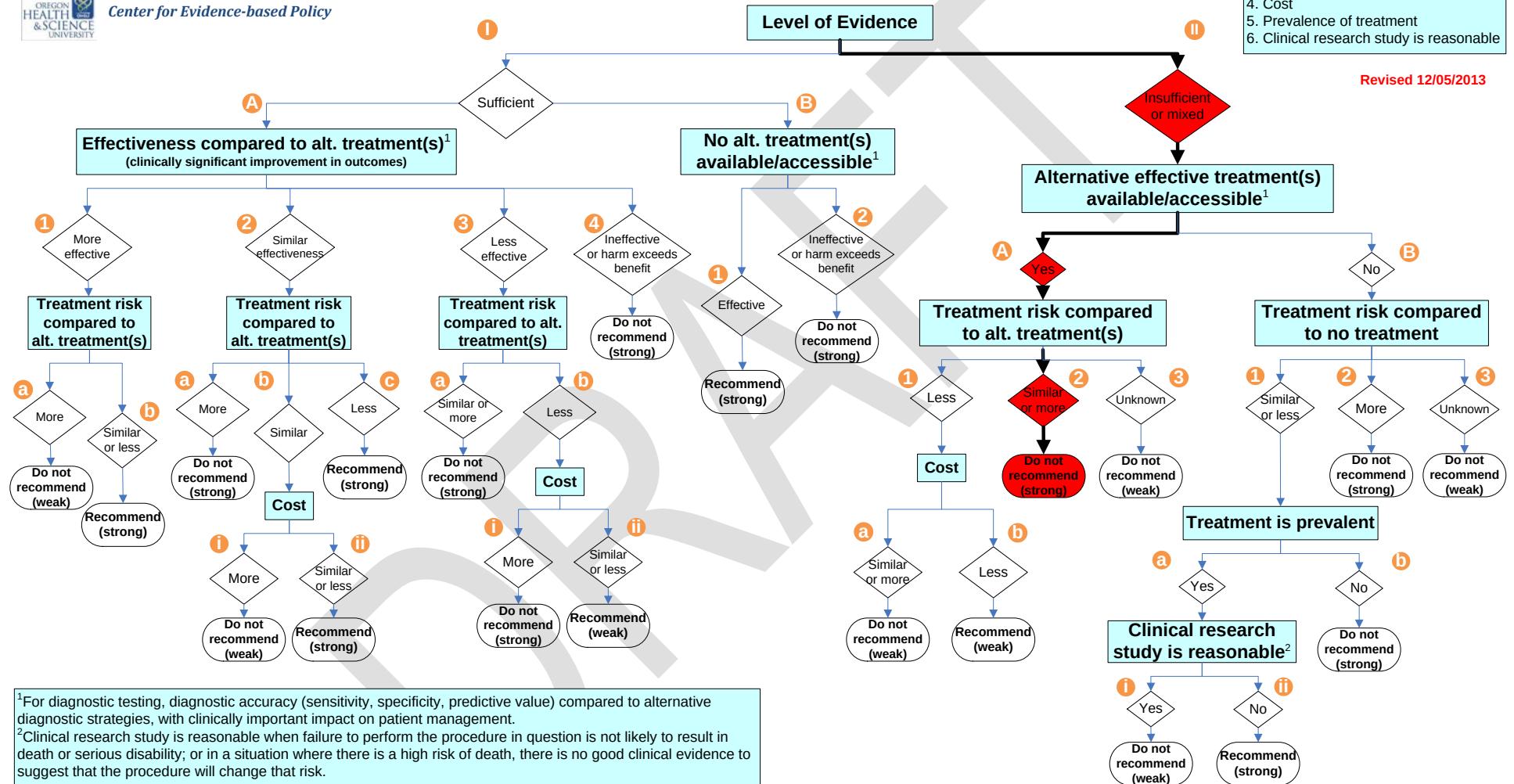
HERC Guidance Development Framework

Refer to *HERC Guidance Development Framework Principles* for additional considerations

Decision Point Priorities

1. Level of evidence
2. Effectiveness & alternative treatments
3. Harms and risk
4. Cost
5. Prevalence of treatment
6. Clinical research study is reasonable

Revised 12/05/2013



¹For diagnostic testing, diagnostic accuracy (sensitivity, specificity, predictive value) compared to alternative diagnostic strategies, with clinically important impact on patient management.

²Clinical research study is reasonable when failure to perform the procedure in question is not likely to result in death or serious disability; or in a situation where there is a high risk of death, there is no good clinical evidence to suggest that the procedure will change that risk.

HERC Coverage Guidance – Nuclear Cardiac Imaging Disposition of Public Comments

Table of Contents

Commenters.....	1
Public Comments	2
References Provided by Commenters	8

Commenters

Identification	Stakeholder
A	Medical Imaging and Technology Alliance, Arlington, VA [Submitted July 23, 2014]
B	Oregon Chapter of the American College of Cardiology, the American College of Cardiology, the American Society of Nuclear Cardiology, the Society of Nuclear Medicine and Molecular Imaging, and the Society of Cardiovascular Computed Tomography [Submitted July 24, 2014]

HERC Coverage Guidance – Nuclear Cardiac Imaging Disposition of Public Comments

Public Comments

Ident.	#	Comment	Disposition
A	1	The Medical Imaging and Technology Alliance (MITA) is pleased to submit comments on the Oregon Health Evidence Review Commission (HERC) draft guidance for Nuclear Cardiac Imaging posted on June 25, 2014. As the leading trade association representing medical imaging, radiotherapy, and radiopharmaceutical manufacturers, MITA has in-depth knowledge of the significant benefits to the health of Americans that medical imaging and radiotherapy provide.	Thank you for taking the time to comment and providing this background.
A	2	Myocardial perfusion imaging is an important and established technology which has provided great benefit to patients suffering from coronary artery disease (CAD) for decades. By using a radiopharmaceutical to pinpoint perfusion defects, cardiologists/doctors are able to identify areas of blockages in the coronary arteries and to determine severity of disease.	EbGS is aware of the value of myocardial perfusion imaging (MPI).
A	3	Because the value of myocardial perfusion imaging is widely acknowledged, we are concerned that the limitations introduced by the draft guidance would negatively affect access to this timely and important modality for your beneficiaries. Instead, we urge you to adopt the established Appropriate Use Criteria (AUC) for Cardiac Radionuclide Imaging (attachment 1). This document represents years of extensive and careful consideration of numerous studies representing thousands of patients and the cumulative experience of experts in cardiology. To wit, the AUC Task Force comprises expertise from the American Society of Nuclear Cardiology, the American College of Cardiology, the American College of Radiology, the American Heart Association, the American Society of Echocardiography, the Society of Cardiovascular Computed Tomography, the Society for Cardiovascular Magnetic Resonance, the Society of Nuclear Medicine and has been endorsed by the American College of Emergency Physicians.	Methodology of appropriate use criteria is a combination of literature review and expert opinion. Neither aspect meets HERC standards as an appropriate data source to override ‘trusted source’ systematic reviews such as those provided by AHRQ or WA HTA. Panel members were “asked to refer to the relevant guidelines for a summary of the relevant literature”. Relationships to industry were reported (and extensive) but not managed in any stated way. Indications for MPI are rated by panel members as “appropriate”, “uncertain”, or “inappropriate”. A total of eight clinical scenarios were rated. The following received “appropriate” scores: • <u>Detection of CAD</u>: symptomatic; evaluation of ischemic equivalent (nonacute) if ECG uninterpretable or unable to exercise, or intermediate or high probability of CAD; • <u>Detection of CAD</u>: symptomatic; acute chest pain (5 different scenarios for possible ACS); • <u>Detection of CAD/risk assessment</u>: without ischemic equivalent; asymptomatic if high risk OR new onset heart failure with LV systolic

HERC Coverage Guidance – Nuclear Cardiac Imaging Disposition of Public Comments

Ident.	#	Comment	Disposition
			dysfunction OR ventricular tachycardia OR syncope OR elevated troponin; • Risk assessment with prior test results and/or known chronic stable angina; prior non-invasive evaluation inconclusive OR new/worsening symptoms OR angiography with coronary stenosis of uncertain significance OR asymptomatic with coronary calcium score 100 to 400 and high risk, or >400 OR duke treadmill score intermediate or high risk; • <u>Risk assessment</u>: Preoperative evaluation for vascular surgery; • <u>Risk assessment</u>: within 3 months of an ACS; • <u>Risk assessment</u>: postrevascularization; • Assessment of viability/ischemia in ischemic cardiomyopathy; and • Evaluation of ventricular function. The following comment is made pertaining to other imaging modalities: “technical panel members were asked to rate indications for cardiac RNI in a manner independent and irrespective of the prior published ACCF ratings for SPECT MPI (1) as well as the prior ACCF ratings for similar diagnostic stress imaging modalities, such as stress echocardiography (2), cardiac computed tomography, or cardiac magnetic resonance (3). Given the iterative nature of the process, readers are counseled not to compare too closely individual appropriate use ratings among modalities rated at different times over the past 2 years.”
A	4	Finally, each patient is unique and, by following the AUC instead of more rigid criteria, decision-making would rightly take place between the physician and the patient to determine the best path for the individual patient’s care, which may or may not include nuclear cardiac imaging at a particular decision point.	EbGS is aware of the uniqueness of patients, but that does not obviate the need for development of coverage policy.

HERC Coverage Guidance – Nuclear Cardiac Imaging Disposition of Public Comments

Ident.	#	Comment	Disposition
A	5	In consideration of the established Appropriate Use Criteria for nuclear cardiac imaging, we strongly recommend that you reconsider your draft guidance and instead follow the AUC as developed by medical professional societies.	The AUC criteria do not compare MPI to other cardiac imaging technologies. The evidence reviewed in the guidance document that compares stress ECHO and MPI indicates equivalence of the two procedures in most circumstances, and stress ECHO is less costly and is associated with less harm.
B	1	The Oregon Chapter of the American College of Cardiology, the American College of Cardiology, the American Society of Nuclear Cardiology, the Society of Nuclear Medicine and Molecular Imaging and the Society of Cardiovascular Computed Tomography appreciate the opportunity to provide comments on the Health Evidence Review Commission's (HERC) Coverage Guidance on Nuclear Cardiac Imaging (NCI).	Thank you for taking the time to comment.
B	2	We strongly disagree with the HERC's draft coverage recommendations for NCI as well its attempting to define a selection algorithm between two appropriate and indicated tests. We do not feel that the algorithm will improve upon the judgment of the ordering physician, lower costs or improve outcomes.	The algorithm (Guidance Development Framework) is a tool to assist the HERC with development of coverage recommendations. It does not apply to individual patients or specific clinical situations.
B	3	We strongly urge the HERC to adopt the recommendations finalized last year by the Washington State Health Care Authority Health Technology Assessment Program on this matter.	The WA HTA decision is as follows: Cardiac Nuclear Imaging is a covered benefit with conditions including: - SPECT - Covered for patients with symptoms of myocardial ischemia (symptomatic) who are: • At high risk of coronary artery disease (CAD), or • At low to intermediate risk of CAD, and - o Have abnormal/indeterminate exercise treadmill test (ETT), or - o Unable to perform ETT, or - o ECG abnormality that prevents accurate interpretation of ETT - For patients with known CAD, monitoring: • Changes in symptoms - PET - Covered under the same conditions as SPECT when: • SPECT is not technically feasible; or

HERC Coverage Guidance – Nuclear Cardiac Imaging Disposition of Public Comments

Ident.	#	Comment	Disposition
			- SPECT is inconclusive Non-Covered Indicators - Cardiac Nuclear Imaging is not a covered benefit for: - Asymptomatic patients* - Patients with known CAD and no changes in symptoms *Does not apply to pre-operative evaluation of patients undergoing high-risk non-cardiac surgery or patients who have undergone cardiac transplant.
B	4	First, we would like to highlight the paper, Appropriate Use Criteria for Cardiac Radionuclide Imaging, published in JACC Vol. 53, No 23, 2009 (June 9, 2009, pages 2201-29), a copy of which is enclosed. The paper covers SPECT myocardial perfusion imaging (MPI) and cardiac PET, and reviews appropriate use of imaging, defined by specific clinical indication (Tables 1-8, Pages 2207-2210). The appropriate use criteria (AUC) describe a rigorous process of literature review and decision making, are endorsed by eight different specialty societies, and have become widely accepted by the medical community and third-party payers. This paper should answer the majority of the commission's concerns and questions about imaging selection for patients. Of note, the terms Appropriate, Uncertain, and Inappropriate have been replaced in new guidelines with the terms Appropriate Care, May be Appropriate Care, and Rarely Appropriate Care (Hendel 2013).	See comment #A3 and #A5
B	5	<u>Sensitivity and Specificity</u> There are a large number of papers which look at MPI accuracy, and the “gold standard” for comparison is most commonly invasive coronary angiography. Generally, the sensitivity and specificity of MPI is in the range of 80-90%. Predictive value, of course, is based upon the pretest prevalence of the disease in the population studied. Though comparative “head to head” studies against stress echo are few, most indicate a superior sensitivity and specificity for MPI, although not all comparisons reach statistical significance. Since sensitivity and specificity of both echo and MPI technologies are quite good, relative differences (favoring MPI) tend to be modest. Subsets where MPI compares particularly favorably to stress echo include patients with single vessel coronary disease, patients who are obese, patients who are unable to exercise or achieve target heart rate, patients with pacemakers or implantable defibrillators or left bundle branch block, patients with prior myocardial infarction or other known coronary or other heart disease, patients with resting ECG abnormalities, and patients with otherwise poor echo image quality. Many patients in	EbGS agrees that sensitivity and specificity of both ECHO and MPI technologies are good, and that relative differences are modest. However, ECHO is not associated with harms of radiation and is substantially less expensive. EbGS agrees that there are certain subpopulations in which MPI compares favorably to ECHO, and these are accounted for in the coverage guidance recommendation: “patients for whom stress imaging is required and stress ECHO is contraindicated or would provide suboptimal imaging* *i.e. pre-existing cardiomyopathy or regional wall motion

HERC Coverage Guidance – Nuclear Cardiac Imaging Disposition of Public Comments

Ident.	#	Comment	Disposition
		these subsets cannot be identified prospectively.	abnormalities, left bundle branch block, paced rhythm, unsuitable acoustic windows due to body habitus, inability to utilize dobutamine in a setting where exercise is not possible or when the target workload is not achievable.”
B	6	- Head-to-head comparisons of MPI versus stress echo in the literature tend to include only patients who have a clinical indication to undergo invasive coronary angiography. In these studies, invasive angiography determines which of the noninvasive tests is “correct”. This approach has several limitations. - First, patients already selected for invasive angiography on clinical grounds have a much higher likelihood of obstructive coronary disease than those undergoing testing in an office setting. Patients with a negative study most often are reassured in practice, and are then seldom referred for angiography. A better outcome measure for patients with a negative test might be “percent event free survival at 3 years,” though this is an exceedingly difficult assessment to make in patients who will not be having regular cardiac follow up.	EbGS agrees that a better outcome measure is event-free survival. Indeed, one study included in the WA HTA review compared the ability of SPECT and stress ECHO to predict major adverse coronary events (MACE) after 3.7 years, and found no significant difference between the two when evaluating the area under the ROC curve. Another study assessed the prognostic value of stress ECHO and SPECT in patients who were followed for 7.3 years and found no statistical differences between the tests. Three additional cohort studies also did not find significant differences in the prognostic ability of SPECT and ECHO.
	7	Second, invasive angiography is an anatomic test, whereas MPI and stress echo are physiologic tests. Angiography will often identify a stenosis in the range of 50-70%, which is considered equivocal on anatomic grounds. Multiple studies have shown that MPI is more likely to demonstrate ischemia in such territories, as compared to stress echo. Though fractional flow reserve assessment can be done at the time of angiography, and does provide a functional assessment of whether such lesions are flow limiting, currently we are not aware of any studies which have utilized this technology in comparison to the results obtained with stress echo or MPI.	Ability to demonstrate ischemia in an area of known stenosis was not specifically addressed in the WA HTA report. However, whether or not this ultimately results in improved patient outcomes is unclear. While commenter states that multiple studies have shown that MPI is more likely to demonstrate ischemia in such territories than stress ECHO, no citations are provided, and no such comparative evidence is presented in the source report.
	8	<u>Frequency of incidental findings (outside the heart)</u> Research indicates that the incidence of noncardiac findings on nuclear MPI is about 2.5% (Williams 2003). Detection of breast cancer, lung cancer, and abnormalities in the thyroid are not uncommon and may result in early detection of disease and subsequent life-saving treatment.	Studies reported in the WA HTA report cite a very low rate of extracardiac findings (0-0.3%) for SPECT. While incidental findings can occasionally result in early detection of disease, more often they result in unnecessary additional testing, increased anxiety and potential harms.
	9	<u>Relevant Outcome Data</u> First, research shows that major adverse cardiac events are better predicted by SPECT than by visually-analyzed coronary angiograms. In addition, there is no incremental prognostic	The WA HTA report does not address the comparison of SPECT with invasive angiography, or the additional value that SPECT may provide.

HERC Coverage Guidance – Nuclear Cardiac Imaging Disposition of Public Comments

Ident.	#	Comment	Disposition
		value from visually-analyzed coronary angiograms over SPECT perfusion data (Iskandrian 1993). • Second, in 2002 a study demonstrated that “ischemia guided angiography” using SPECT was more cost effective than routine angiography in patients with stable chest pain (Shaw 2002). • Further, Hachamovitch et al. (2003) published research that showed patients with a lesser degree of reversible ischemia on SPECT perfusion imaging had higher survival rates with medical therapy than with revascularization. However, those with a greater degree of ischemia were likely to benefit from invasive procedures. A subsequent study demonstrated that women who had a “false positive” noninvasive study and thus had coronary angiography showing “no significant” stenosis often had endothelial dysfunction documented in the catheterization lab. • Finally, the FAME trial affirmed the prognostic value of physiologic assessment to document ischemia before performing revascularization (Tonino 2009).	Iskandrian 1993 is a case series of 316 medically treated patients with CAD who had undergone both angiography and stress MPI who were followed for a mean of 28 months. Cox regression models were used to determine the prognostic values of clinical, exercise, MPI and angiography data in predicting cardiac death or MI. Authors report that extent of perfusion abnormality had the higher chi-square value of any of the variables (gender, exercise work, cath data). Shaw 2002 was a retrospective cohort study that compared direct referral for angiography to stress myocardial perfusion imaging in women with stable chest pain. Cost per patient was calculated for both strategies and found to be \$2490 for angiography and \$1587 for MPI in low risk patients; for high risk patients, the costs were \$3687 and \$2585 respectively. Hachomovitch 2003 was a retrospective cohort study of patients without prior MI or revascularization who received SPECT. Patients were assigned to groups based on treatment received at 60 days after SPECT (revascularization or medical therapy). Patients were then followed approx. 2 years and outcomes of cardiac death and all-cause death were calculated. Cardiac death was higher in the revascularization group (2.8% vs 1.3%). However, authors state that “with increasing amounts of inducible ischemia, mortality rates progressively increased in patients undergoing medical therapy”, and that in patients with > 20% myocardium ischemic, revascularization had a lower cardiac death rate than medical therapy. Tonino 2009 evaluated fractional flow reserve calculated

HERC Coverage Guidance – Nuclear Cardiac Imaging Disposition of Public Comments

Ident.	#	Comment	Disposition
			from angiography; MPI was not utilized. None of these studies would meet the inclusion criteria of the WA HTA review, which included only RCTs or observational studies comparing alternative testing strategies; none of them compare the diagnostic value of ECHO with SPECT. There is evidence demonstrating that the prognostic value of stress echo and MPI are comparable.
	10	<u>Summary</u> Stress echo, MPI and cardiac PET have AUC which have been rigorously developed, endorsed by multiple specialty societies, and utilized by the medical community successfully for years. There are certainly circumstances where patients referred for imaging may meet AUC for both stress echo and MPI. Selecting the appropriate test in these circumstances is generally best left to the ordering physician, using the factors described above, along with their knowledge of the individual patient.	Selection of the appropriate test is always left to the ordering physician; that does not obviate the need for coverage policy development.
	11	Again, we support the HERC adopting the NCI guidance based on the AUC and the Washington State Health Care Authority Technology Assessment. Thank you for this opportunity to comment on the Health Evidence Review Commission's Coverage Guidance on Nuclear Cardiac Imaging.	The AUC criteria do not compare MPI to other cardiac imaging technologies. The evidence reviewed in the guidance document that compares stress ECHO and MPI indicates equivalence of the two procedures in most circumstances, and stress ECHO is less costly and is associated with less harm.

HERC Coverage Guidance – Nuclear Cardiac Imaging Disposition of Public Comments

References Provided by Commenters

Commenter	References
A	(1) Hendel, R. C., Berman, D. S., Di Carli, M. F., Heidenreich, P. A., Henkin, R. E., Pellikka, P. A., Williams, K. A., et al. (2009). ACCF/ASNC/ACR/AHA/ASE/SCCT/SCMR/SNM 2009 Appropriate Use Criteria for Cardiac Radionuclide Imaging. A Report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, the American Society of Nuclear Cardiology, the American College of Radiology, the American Heart Association, the American Society of Echocardiography, the Society of Cardiovascular Computed Tomography, the Society for Cardiovascular Magnetic Resonance, and the Society of Nuclear Medicine Endorsed by the American <i>Journal of the American College of Cardiology</i> , 53(23), 2201-2229.
B	(1) Hendel, R. C., Berman, D. S., Di Carli, M. F., Heidenreich, P. A., Henkin, R. E., Pellikka, P. A., Williams, K. A., et al. (2009). ACCF/ASNC/ACR/AHA/ASE/SCCT/SCMR/SNM 2009 Appropriate Use Criteria for Cardiac Radionuclide Imaging. A Report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, the American Society of Nuclear Cardiology, the American College of Radiology, the American Heart Association, the American Society of Echocardiography, the Society of Cardiovascular Computed Tomography, the Society for Cardiovascular Magnetic Resonance, and the Society of Nuclear Medicine Endorsed by the American<i>Journal of the American College of Cardiology</i>, 53(23), 2201-2229. (2) Hendel, R. C., Patel, M. R., Allen, J. M., Min, J. K., Shaw, L. J., Wolk, M. J., et al. (2013). Appropriate use of cardiovascular technology: 2013 ACCF appropriate use criteria methodology update: a report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force. <i>Journal of the American College of Cardiology</i>, 61(12), 1305-1317. (3) Williams, K. A., Hill, K. A., & Sheridan, C. M. (2003). Noncardiac findings on dual-isotope myocardial perfusion SPECT. <i>Journal of nuclear cardiology</i>, 10(4), 395-402. (4) Iskandrian, A. S., Chae, S. C., Heo, J., Stanberry, C. D., & Wasserleben, V. (1993). Independent and incremental prognostic value of exercise single-photon emission computed tomographic (SPECT) thallium imaging in coronary artery disease. <i>Journal of the American College of Cardiology</i>, 22(3), 665-670. (5) Shaw, L. J., Heller, G. V., Travin, M. I., Lauer, M., Marwick, T., Hachamovitch, R., et al. (1999). Cost analysis of diagnostic testing for coronary artery disease in women with stable chest pain. <i>Journal of Nuclear Cardiology</i>, 6(6), 559-569. (6) Hachamovitch, R., Hayes, S. W., Friedman, J. D., Cohen, I., & Berman, D. S. (2003). Comparison of the short-term survival benefit associated with revascularization compared with medical therapy in patients with no prior coronary artery disease undergoing stress myocardial perfusion single photon emission computed tomography. <i>Circulation</i>, 107(23), 2900-2907. (7) Tonino, P. A., De Bruyne, B., Pijls, N. H., Siebert, U., Ikeno, F., vant Veer, M., et al. (2009). Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. <i>New England Journal of Medicine</i>, 360(3), 213-224.

Section 6.0

Previously Discussed Items

Treatment of obstructive sleep apnea in children with craniofacial anomalies

Question: Treatment of OSA (sleep apnea) in children with craniofacial anomalies

Question source: HERC Staff

Issue: In May 2014 VbBS and HERC approved a modified guideline on the treatment of sleep apnea based on a review in adults. As such, surgical codes for the treatment of sleep apnea in adults were removed from the higher sleep apnea line 210 as there is insufficient evidence of effectiveness. Tonsillectomy and adenoidectomy codes were left on Line 210 for children alone.

However, children with craniofacial anomalies may need these surgical procedures and the removal of these codes may cause significant challenge when trying to get procedures covered.

From Anna Kuang, Craniofacial plastic surgeon at OHSU

We have a significant subset of craniofacial patients who need jaw surgery (distraction or formal jaw surgery) as means to obviate need for tracheostomy and or improve airway. Mandibular osteotomies the only surgical option besides tracheostomy (T&As helpful for managing them post op).

Commonly affected kids:

Any kid with a craniofacial syndrome may have associated micrognathia and glossoptosis WITH or WITHOUT cleft palate. And they can have midface hypoplasia that need midface advancement or distraction also WITH or WITHOUT cleft palate.

Kids with skeletal dysplasia have associated micrognathia and glossoptosis

Pierre Robin Sequence diagnosis is micrognathia, glossoptosis, WITH or WITHOUT cleft palate

Cleft Lip/Palate and Craniofacial syndromic kids have congenital nasal obstruction, deviated septum, and or hypoplastic noses and need opens septorhinoplasty and other nasal surgeries.

Diagnosis of OSA in pediatric population in difficult (multiple sleep studies, airway evals, CT/MRI's, repeat visits, multiple disciplines) and time consuming. Once the diagnosis is made, we often have to proceed quickly with surgical intervention. Only a small percentage of kids have CPAP as an option due to age and age/diagnosis compliance issues. Delay in approval or surgical intervention procedures is associated with increased morbidity.

GUIDELINE NOTE 27, TREATMENT OF SLEEP APNEA IN ADULTS

Line 210

CPAP is covered initially when all of the following conditions are met:

Treatment of OSA (sleep apnea) in children with craniofacial anomalies,

Issue #780

Page 1

Treatment of obstructive sleep apnea in children with craniofacial anomalies

- 12 week 'trial' period to determine benefit. This period is covered if apnea-hypopnea index (AHI) or respiratory disturbance index (RDI) is greater than or equal to 15 events per hour; or if between 5 and 14 events with additional symptoms including one or more of the following:
 - o excessive daytime sleepiness defined as either an Epworth Sleepiness Scale score >10 or daytime sleepiness interfering with ADLs, that is not attributable to another modifiable sedating condition (e.g. narcotic dependence), or
 - o documented hypertension, or
 - o ischemic heart disease, or
 - o history of stroke;
- Providers must provide education to patients and caregivers prior to use of CPAP machine to ensure proper use; and
- Positive diagnosis through polysomnogram (PSG) or Home Sleep Test (HST).

CPAP coverage subsequent to the initial 12 weeks is based on documented patient tolerance, compliance, and clinical benefit. Compliance (adherence to therapy) is defined as use of CPAP for at least four hours per night on 70% of the nights during a consecutive 30 day period.

Mandibular advancement devices (oral appliances) are covered for those for whom CPAP fails or is contraindicated.

Surgery for sleep apnea in adults is not covered (due to lack of evidence of efficacy). Tonsillectomy and adenoidectomy codes are included on this line only for children who meet criteria according to Guideline Note 118 OBSTRUCTIVE SLEEP APNEA DIAGNOSIS AND TREATMENT IN CHILDREN.

GUIDELINE NOTE 118, OBSTRUCTIVE SLEEP APNEA DIAGNOSIS AND TREATMENT FOR CHILDREN (*being reviewed this meeting*)

Line 206

Obstructive sleep apnea (OSA) in children (18 or younger) must be diagnosed by

1. nocturnal polysomnography with an AHI >5 episodes/h or AHI >1 episodes/h with history and exam consistent with OSA, OR
2. nocturnal pulse oximetry with 3 or more SpO₂ drops <90% and 3 or more clusters of desaturation events, or alternatives desaturation (>3%) index >3.5 episodes/h, OR
3. use of a validated questionnaire (such as the Pediatric Sleep Questionnaire or OSA 18), OR
4. consultation with a sleep medicine specialist.

Polysomnography and/or consultation with a sleep medicine specialist to support the diagnosis of OSA and/or to identify perioperative risk is recommended for

1. high risk children (i.e. children with cranio-facial abnormalities, neuromuscular disorders, Down syndrome, etc.)

Treatment of obstructive sleep apnea in children with craniofacial anomalies

2. children with equivocal indications for adenotonsillectomy (such as discordance between tonsillar size on physical examination and the reported severity of sleep-disordered breathing),
3. children younger than three years of age

Adenotonsillectomy is an appropriate first line treatment for children with OSA. Weight loss is recommended in addition to other therapy in patients who are overweight or obese. [Adenoidectomy without tonsillectomy is only covered when a child with OSA has previously had a tonsillectomy, when tonsillectomy is contraindicated, or when tonsillar hypertrophy is not present. More complex surgical treatments are only included on this line for children with craniofacial anomalies.](#)

Intranasal corticosteroids are an option for children with mild OSA in whom adenotonsillectomy is contraindicated or for mild postoperative OSA.

CPAP is covered for a 3 month trial for children through age 18 who have

1. undergone surgery or are not candidates for surgery, AND
2. have documented residual sleep apnea symptoms (sleep disruption and/or significant desaturations) with residual daytime symptoms (daytime sleepiness or behavior problems)

CPAP will be covered for children through age 18 on an ongoing basis if:

- There is documentation of improvement in sleep disruption and daytime sleepiness and behavior problems with CPAP use
- Annual re-evaluation for CPAP demonstrates ongoing clinical benefit and compliance with use, defined as use of CPAP for at least four hours per night on 70% of the nights in a consecutive 30 day period

Code placement completed May 2014 VbBS/HERC meeting

Code	Description	Recommendation
21193	Reconstruction of mandibular rami, horizontal, vertical, C, or L osteotomy; without bone graft	Remove from 210. (Remains only on line 646)
21194	Reconstruction of mandibular rami, horizontal, vertical, C, or L osteotomy; with bone graft (includes obtaining graft)	Remove from 210. (Remains only on line 646)
21195	Reconstruction of mandibular rami and/or body, sagittal split; without internal rigid fixation	Remove from 210. (Remains only on line 646)
21196	Reconstruction of mandibular rami and/or body, sagittal split; with internal rigid fixation	Remove from 210. (Remains only on line 646)
21198	Osteotomy, mandible, segmental;	Remove from 210. (Remains only on line 646)
21199	Osteotomy, mandible, segmental; with genioglossus advancement	Remove from 210. (Remains only on line 646)
21206	Osteotomy, maxilla, segmental (eg, Wassmund or Schuchard)	Remove from 210. (Remains only on line 646)
21208	Osteoplasty, facial bones; augmentation (autograft, allograft, or prosthetic implant)	Remove from 210. (Remains only on line 646)

Treatment of obstructive sleep apnea in children with craniofacial anomalies

Code	Description	Recommendation
21209	Osteoplasty, facial bones; reduction	Remove from 210. (Remains only on line 646)
21210	Graft, bone; nasal, maxillary or malar areas (includes obtaining graft)	Remove from 210. (Remains on lines 242,612,673).
21215	Graft, bone; mandible (includes obtaining graft)	Remove from 210. (Remains on lines 242,612,673).
21230	Graft; rib cartilage, autogenous, to face, chin, nose or ear (includes obtaining graft)	Remove from 210. (Remains on line 673).
21235	Graft; ear cartilage, autogenous, to nose or ear (includes obtaining graft)	Remove from 210. (Remains on 405,673).
30117	Excision or destruction (eg, laser), intranasal lesion; internal approach	Remove from 210. (Remains on multiple other lines)
30140	Submucous resection inferior turbinate, partial or complete, any method	Remove from 210. (Remains on multiple other lines)
30520	Septoplasty or submucous resection, with or without cartilage scoring, contouring or replacement with graft	Remove from 210. (Remains on multiple other lines)
42140	Uvulectomy, excision of uvula	Remove from 210. Place on DMAP Excluded File.
42145	Palatopharyngoplasty (eg, uvulopalatopharyngoplasty, uvulopharyngoplasty)	Remove from 210. Will remain on 71 and 325
42160	Destruction of lesion, palate or uvula (thermal, cryo or chemical)	Remove from 210. (Remains on 656).

HERC Staff Recommendations:

- 1) Add surgical codes in above table back to Line 210 (in addition to the preexisting tonsillectomy and adenoidectomy codes)
- 2) Delete the coding specification from Line 210 as it is no longer necessary given the revised guideline that states surgery for sleep apnea in adults is not covered
~~42299 Unlisted procedure, palate, uvula (use for laser assisted uvulopalatoplasty (LAUP), somnoplasty, palatal implants) does not pair on Line 210 with obstructive sleep apnea in adults.~~
- 3) Modify Guideline Note 27 as follows

GUIDELINE NOTE 27, TREATMENT OF SLEEP APNEA IN ADULTS

Line 210

CPAP is covered initially when all of the following conditions are met:

- 12 week 'trial' period to determine benefit. This period is covered if apnea-hypopnea index (AHI) or respiratory disturbance index (RDI) is greater than or equal to 15 events per hour; or if between 5 and 14

Treatment of obstructive sleep apnea in children with craniofacial anomalies

events with additional symptoms including one or more of the following:

- o excessive daytime sleepiness defined as either an Epworth Sleepiness Scale score >10 or daytime sleepiness interfering with ADLs, that is not attributable to another modifiable sedating condition (e.g. narcotic dependence), or
 - o documented hypertension, or
 - o ischemic heart disease, or
 - o history of stroke;
- Providers must provide education to patients and caregivers prior to use of CPAP machine to ensure proper use; and
- Positive diagnosis through polysomnogram (PSG) or Home Sleep Test (HST).

CPAP coverage subsequent to the initial 12 weeks is based on documented patient tolerance, compliance, and clinical benefit. Compliance (adherence to therapy) is defined as use of CPAP for at least four hours per night on 70% of the nights during a consecutive 30 day period.

Mandibular advancement devices (oral appliances) are covered for those for whom CPAP fails or is contraindicated.

Surgery for sleep apnea in adults is not covered (due to lack of evidence of efficacy). ~~Tonsillectomy and adenoidectomy~~ [Surgical](#) codes are included on this line only for children who meet criteria according to Guideline Note 118 OBSTRUCTIVE SLEEP APNEA DIAGNOSIS AND TREATMENT IN CHILDREN.

- 4) Modify Guideline Note 118 as follows:

GUIDELINE NOTE 118, OBSTRUCTIVE SLEEP APNEA DIAGNOSIS AND TREATMENT FOR CHILDREN *(being reviewed this meeting)*

Line 206

Obstructive sleep apnea (OSA) in children (18 or younger) must be diagnosed by

1. nocturnal polysomnography with an AHI >5 episodes/h or AHI >1 episodes/h with history and exam consistent with OSA, OR
2. nocturnal pulse oximetry with 3 or more SpO₂ drops <90% and 3 or more clusters of desaturation events, or alternatives desaturation (>3%) index >3.5 episodes/h, OR
3. use of a validated questionnaire (such as the Pediatric Sleep Questionnaire or OSA 18), OR
4. consultation with a sleep medicine specialist.

Polysomnography and/or consultation with a sleep medicine specialist to support the diagnosis of OSA and/or to identify perioperative risk is recommended for

1. high risk children (i.e. children with cranio-facial abnormalities, neuromuscular disorders, Down syndrome, etc.)

Treatment of obstructive sleep apnea in children with craniofacial anomalies

2. children with equivocal indications for adenotonsillectomy (such as discordance between tonsillar size on physical examination and the reported severity of sleep-disordered breathing),
3. children younger than three years of age

Adenotonsillectomy is an appropriate first line treatment for children with OSA. Weight loss is recommended in addition to other therapy in patients who are overweight or obese. [Adenoidectomy without tonsillectomy is only covered when a child with OSA has previously had a tonsillectomy, when tonsillectomy is contraindicated, or when tonsillar hypertrophy is not present. More complex surgical treatments are only included on this line for those with craniofacial anomalies.](#)

Intranasal corticosteroids are an option for children with mild OSA in whom adenotonsillectomy is contraindicated or for mild postoperative OSA.

CPAP is covered for a 3 month trial for children through age 18 who have

1. undergone surgery or are not candidates for surgery, AND
2. have documented residual sleep apnea symptoms (sleep disruption and/or significant desaturations) with residual daytime symptoms (daytime sleepiness or behavior problems)

CPAP will be covered for children through age 18 on an ongoing basis if:

- There is documentation of improvement in sleep disruption and daytime sleepiness and behavior problems with CPAP use
- Annual re-evaluation for CPAP demonstrates ongoing clinical benefit and compliance with use, defined as use of CPAP for at least four hours per night on 70% of the nights in a consecutive 30 day period

Section 7.0

Guidelines

Adenoidectomy

Question: When is adenoidectomy alone indicated and/or effective?

Question source: Tracy Muday, MD, OHP Medical Director

Issue:

From Dr. Muday:

We have lots of requests for adenoidectomy alone for young kids, usually under 4, sometimes under 2. My understanding of why they are doing the adenoidectomy alone is related to the risk of adenotonsillectomy in the younger group, and they feel that the adenoidectomy alone is likely to improve the kid's symptoms.

Truthfully, my 3 ENTs have a gestalt of kids that they feel need tubes or adenoidectomy. The diagnosis that they use for this adjusts itself depending on what is covered. When OME was above the line, all the tubes were done with a dx of OME. When OME went below the line, they started paying attention to who had recurrent AOM, and that became the diagnosis. When we put limits on adenoidectomy with tubes for OME, they started coding for OSA and sinusitis

Adenoidectomy for chronic otitis media

Adenoidectomy for treatment of chronic otitis media was extensively reviewed in 2012 and guideline revisions made to clarify when this procedure is appropriate. Adenoidectomy was reaffirmed as a treatment for possible children with chronic otitis media receiving their second set of tubes in the Chronic Otitis Media guideline.

Adenoidectomy for chronic sinusitis

Adenoidectomy for chronic sinusitis was last reviewed in 2008. Some evidence was found that younger children (younger than age 12) may benefit for adenoidectomy for chronic sinusitis, but overall the body of evidence was scanty. Adenoidectomy was added to the chronic sinusitis line based mainly on expert opinion. It was noted that "If surgery is indicated, evidence shows that adenoidectomy should be performed before endoscopic sinus surgery (which is currently covered)." Adenoidectomy was added to the chronic sinusitis line for children younger than age 12 (42830, 42835).

Adenoidectomy for OSA

The obstructive sleep apnea guideline (GN 118) was extensively reviewed in 2013, but adenoidectomy as a single procedure was not reviewed. Evidence presented during this review focused on adenotonsillectomy.

Adenoidectomy

Current List placement:

CPT	Code Description	Current Lines
42820	Tonsillectomy and adenoidectomy; younger than age 12	47 CONGENITAL AIRWAY OBSTRUCTION WITH OR WITHOUT CLEFT PALATE 51 DEEP ABSCESES, INCLUDING APPENDICITIS AND PERIORBITAL ABSCESS; INTESTINAL PERFORATION 68 CONGENITAL ANOMALIES OF UPPER ALIMENTARY TRACT, EXCLUDING TONGUE 206 SLEEP APNEA AND NARCOLEPSY 370 STREPTOCOCCAL SORE THROAT AND SCARLET FEVER; VINCENT'S DISEASE; ULCER OF TONSIL; UNILATERAL HYPERTROPHY OF TONSIL 556 CHRONIC DISEASE OF TONSILS AND ADENOIDS
42821	Tonsillectomy and adenoidectomy; age 12 or over	47,68,51,206,370,556
42830	Adenoidectomy, primary; younger than age 12	47,51,206, 470 CHRONIC SINUSITIS 481 CHRONIC OTITIS MEDIA 556
42831	Adenoidectomy, primary; age 12 or over	47,51,206,481,556
42835	Adenoidectomy, secondary; younger than age 12	47,51,206,470,481,556
42836	Adenoidectomy, secondary; age 12 or over	47,51,206,481,556

GUIDELINE NOTE 35, SINUS SURGERY

Lines 366,470

Sinus surgery indicated in the following circumstances:

A) 4 or more episodes of acute rhinosinusitis in one year

OR

B) Failure of medical therapy of chronic sinusitis including all of the following:

- Several courses of antibiotics AND
- Trial of inhaled and/or oral steroids AND
- Allergy assessment and treatment when indicated AND
- One or more of the following:
 - Findings of obstruction of active infection on CT scan
 - Symptomatic mucocele
 - Negative CT scan but significant disease found on nasal endoscopy

OR

C) Nasal polyposis causing or contributing to sinusitis

OR

D) Complications of sinusitis including subperiosteal or orbital abscess, Pott's puffy tumor, brain abscess or meningitis

OR

E) Invasive or allergic fungal sinusitis

OR

F) Tumor of nasal cavity or sinuses

OR

G) CSF rhinorrhea

Adenoidectomy

GUIDELINE NOTE 36, ADENOTONSILLECTOMY FOR INDICATIONS OTHER THAN OBSTRUCTIVE SLEEP APNEA

Lines 47,51,370,556

Tonsillectomy/adenotonsillectomy is an appropriate treatment for patients with:

- A) Five documented attacks of strep tonsillitis in a year or 3 documented attacks of strep tonsillitis in each of two consecutive years where an attack is considered a positive culture/screen and where an appropriate course of antibiotic therapy has been completed;
- B) Peritonsillar abscess requiring surgical drainage; or,
- C) Unilateral tonsillar hypertrophy in adults; unilateral tonsillar hypertrophy in children with other symptoms suggestive of malignancy.

See Guideline Note 118 for diagnosis and treatment of obstructive sleep apnea in children.

GUIDELINE NOTE 51, CHRONIC OTITIS MEDIA WITH EFFUSION

Lines 317,481

Antibiotic and other medication therapy (including antihistamines, decongestants, and nasal steroids) are not indicated for children with chronic otitis media with effusion (OME) (without another appropriate diagnosis).

Patients with specific higher risk conditions (including craniofacial anomalies, Down's syndrome, and cleft palate, or documented speech and language delay) along with hearing loss and chronic otitis media with effusion are intended to be included on Line 317. Otherwise hearing loss associated with chronic otitis media with effusion (without those specific higher risk conditions) is only included on Line 481.

For coverage to be considered on either Line 317 or Line 481, there should be a 3 to 6 month watchful waiting period after diagnosis of otitis media with effusion, and if documented hearing loss is greater than or equal to 25dB in the better hearing ear, tympanostomy surgery may be indicated, given short- but not long- term improvement in hearing. Formal audiometry is indicated for children with chronic OME present for 3 months or longer. Children with language delay, learning problems, or significant hearing loss should have hearing testing upon diagnosis. Children with chronic OME who are not at risk for language delay (such as those with hearing loss <25dB in the better hearing ear) or developmental delay (should be reexamined at 3- to 6-month intervals until the effusion is no longer present, significant hearing loss is identified, or structural abnormalities of the eardrum or middle ear are suspected).

Adenoidectomy is not indicated at the time of first pressure equalization tube insertion. It may be indicated in children over 3 years who are having their second set of tubes.

GUIDELINE NOTE 118, OBSTRUCTIVE SLEEP APNEA DIAGNOSIS AND TREATMENT FOR CHILDREN

Line 206

Obstructive sleep apnea (OSA) in children (18 or younger) must be diagnosed by

Adenoidectomy

1. nocturnal polysomnography with an AHI >5 episodes/h or AHI>1 episodes/h with history and exam consistent with OSA, OR
2. nocturnal pulse oximetry with 3 or more SpO2 drops <90% and 3 or more clusters of desaturation events, or alternatives desaturation (>3%) index >3.5 episodes/h, OR
3. use of a validated questionnaire (such as the Pediatric Sleep Questionnaire or OSA 18), OR
4. consultation with a sleep medicine specialist.

Polysomnography and/or consultation with a sleep medicine specialist to support the diagnosis of OSA and/or to identify perioperative risk is recommended for

1. high risk children (i.e. children with cranio-facial abnormalities, neuromuscular disorders, Down syndrome, etc.)
2. children with equivocal indications for adenotonsillectomy (such as discordance between tonsillar size on physical examination and the reported severity of sleep-disordered breathing),
3. children younger than three years of age

Adenotonsillectomy is an appropriate first line treatment for children with OSA. Weight loss is recommended in addition to other therapy in patients who are overweight or obese.

Intranasal corticosteroids are an option for children with mild OSA in whom adenotonsillectomy is contraindicated or for mild postoperative OSA.

CPAP is covered for a 3 month trial for children through age 18 who have

1. undergone surgery or are not candidates for surgery, AND
2. have documented residual sleep apnea symptoms (sleep disruption and/or significant desaturations) with residual daytime symptoms (daytime sleepiness or behavior problems)

CPAP will be covered for children through age 18 on an ongoing basis if:

- There is documentation of improvement in sleep disruption and daytime sleepiness and behavior problems with CPAP use
- Annual re-evaluation for CPAP demonstrates ongoing clinical benefit and compliance with use, defined as use of CPAP for at least four hours per night on 70% of the nights in a consecutive 30 day period

Adenoidectomy

Evidence—Adenoidectomy for chronic sinusitis

No NICE or SIGN guidances found

AAP and IDSA guidelines on chronic sinusitis in children addressed only diagnostic criteria and antibiotic treatment. Neither addressed surgery

- 1) **van den Aardweg 2010**, Cochrane review of adenoidectomy for chronic nasal symptoms in children
 - a) N=1 study with 76 subjects who underwent adenoidectomy for improvement in episodes of common colds
 - b) The risk differences were non-significant, being 2% (95% CI -18% to 22%) and -11% (95% CI -28% to 7%) after 12 and 24 months, respectively.
 - c) Authors' conclusions: Current evidence regarding the effect of adenoidectomy on recurrent or chronic nasal symptoms or nasal obstruction alone is sparse, inconclusive and has a significant risk of bias.
- 2) **Brietzke 2008**, meta-analysis of adenoidectomy for treatment of pediatric rhinosinusitis
 - a) N=9 studies (10-121 subjects)
 - i) 6 cohort, 4 case-control
 - b) All studies showed that sinusitis symptoms or outcomes improved in half or more patients after adenoidectomy.
 - c) The summary estimate of the proportion of patients who significantly improved after adenoidectomy was 69.3% (95% CI = 56.8—81.7%, $p < 0.001$).
 - d) Conclusion: Adenoidectomy reduces caregiver reported symptoms of chronic rhinosinusitis in a majority of pediatric patients. Given its simplicity, low risk profile, and apparent effectiveness, adenoidectomy should be considered first line therapy for medically refractory, uncomplicated pediatric rhinosinusitis.

Evidence: adenoidectomy for OSA

1) **Shintani 1998**, cohort study

- a. N=13 children who underwent adenoidectomy alone (part of a larger cohort study of 134 children with OSA who underwent various surgical interventions)
 - i. Adenoidectomy was only indicated for children without tonsillar hypertrophy.
- b. Significant improvements of AHI and the lowest SaO₂ were noted 2 months after the treatments in 84.5% (11:13) of children undergoing adenotonsillectomy only and in 75.4% (86:114) of children undergoing adenotonsillectomy
- c. 2 out of the 13 adenoidectomy alone children required repeat surgery within 4-9 months for recurrent OSA compared to 0 of the adenotonsillectomy group

2) **Hensley 2011**, review of adenoidectomy for OSA

- a. N=6 studies included with adenoidectomy for OSA, including Shintani 1998
 - i. 1 cohort study of 206 children who underwent adenoidectomy for OSA. Follow up 3-5 years. Questionnaire results showed 74–87% improvement in symptoms. No PSG data.
 - ii. 1 prospective cohort with 21 children with adenoidectomy, part of larger study with various surgical interventions. Analysis performed for whole group – reduction in mean AHI 11.04→2.14.
 - iii. 3 retrospective cohort studies, looked at rate of re-operation after adenoidectomy. 0.55-29% had subsequent tonsillectomy/ revision adenoidectomy. (N=100-13,005)

Adenoidectomy

Other guidelines

- 1) **AAP practice guideline 2012** for treatment of OSA in children
 - a. Adenotonsillectomy is very effective in treating OSAS. Adenoidectomy or tonsillectomy alone may not be sufficient, because residual lymphoid tissue may contribute to persistent obstruction.

Expert Input—Adenoidectomy for OSA

Derek Lam, MD, OHSU ENT

I'll have to take some time to do a more thorough literature review so I can provide some evidence, but off the top of my head I can tell you that there are definitely patients, especially younger ones, say 12-24 months for whom adenoidectomy alone may absolutely be an appropriate procedure for OSA. Often this may be the primary source of obstruction when the tonsils are small, and there is far less morbidity related to an adenoidectomy than a T&A (less pain, less potential for bleeding, faster recovery). As such, I think it would be a mistake to removal OSA as an indication for adenoidectomy alone. Because that would effectively put ENTs like myself in the uncomfortable position of recommending a T&A in order to get surgery authorized when an adenoidectomy alone would potentially suffice in situations where patients are clearly obstructing and need to have surgery. Admittedly this is a relatively uncommon clinical scenario, but certainly not so uncommon as to warrant removing this as an indicated procedure for OSA.

Dr. Ericka King, MD, OHSU ENT

I agree with Derek that there is certainly a cohort for whom adenoidectomy alone is a reasonable surgical intervention. There is fairly limited data, but I was able to find one study that directly assessed the question: Shintani et al. Int J Pediatr Otorhinolaryngol 1998; 44:51. In this study, they selected children with an AHI of >10. Those with minimal tonsillar hypertrophy underwent adenoidectomy alone, with n=13. In this group, 84.5% demonstrated improvement on postop sleep study (defined as reduction of AHI by at least 1/2, with SpO2 nadir >85%). In the same study, kids that underwent T&A interestingly only had a 75.4% rate of improvement. This is obviously a small study and limited, but one of the few out there. The risks and recovery associated with adenoidectomy alone are much lower than a tonsillectomy, and I feel it is a very reasonable and valuable intervention in selected young children (<3 yrs) with isolated adenoid hypertrophy and history/PSG suggestive of OSA.

Expert Input—adenoidectomy for chronic sinusitis

Dr. Ericka King, MD, OHSU ENT

I disagree that the evidence supporting adenoidectomy is poor; actually it's fairly robust. The role in CRS is likely to be due to biofilm formation/ functioning of the adenoid pad as a bacterial reservoir; there is excellent evidence for this that I don't have time to summarize right now. Clinically, adenoidectomy alone has been shown to result in significant improvement in many studies, which are variably powered. Here are a few:

- Takahashi et al. Acta Otorhinolaryngol Belg 1997; 51:85. This study looked at 78 kids aged 5-7 with sinusitis and COMET; 45 underwent adenoidectomy alone, the other 33 were controls. They found an improvement in symptoms with 56% after adenoidectomy vs 24% of the control group.
- Vanderberg et al. Arch Otolaryngol Head Neck Surg 1997; 123:675. This group looked at 48 kids with CRS, and found complete or near complete symptom resolution in 58%,

Adenoidectomy

some improvement in 21%, and minimal/no improvement in 21% after adenoidectomy alone.

- Ramadan et al. Laryngoscope 2004; 114:2103. The authors studied 202 children with chronic or recurrent acute sinusitis by clinical and CT criteria that failed extensive medical management. They were randomized to endoscopic sinus surgery (ESS) + adenoidectomy, ESS alone, or adenoidectomy alone. These groups showed 87.%, 75%, and 52% improvement rates respectively, with children 6 years and under less likely to improve than older children (60% vs 84% overall).

Maximal medical management generally consists of prolonged antibiotics (3-6 weeks depending on the paper), nasal steroids and saline irrigations. This is a very difficult regimen for many young children, particularly the topical therapies. Additionally, for very good reasons we are trying to avoid CT scanning, particularly in young children. Thus, especially for young kids, I feel it is very reasonable to start with an adenoidectomy instead of a trial of medical therapy.

Summary: Evidence supports use of adenoidectomy for medically refractory chronic sinusitis in children. Based on limited evidence, adenoidectomy alone appears to be effective for treatment of OSA in children without tonsillar hypertrophy.

Adenoidectomy

HERC staff recommendations:

- 1) Continue coverage of adenoidectomy for chronic sinusitis in children
 - a. Clarify GN 35 to indicate when adenoidectomy is indicated as shown below
 - i. Consider whether inhaled or oral steroids are appropriate in young children
- 2) Clarify coverage for adenoidectomy for OSA in children in GN 118 as shown below

GUIDELINE NOTE 35, SINUS SURGERY

Lines 366,470

Sinus surgery ([other than adenoidectomy](#)) is indicated in the following circumstances:

- A) 4 or more episodes of acute rhinosinusitis in one year

OR

- B) Failure of medical therapy of chronic sinusitis including all of the following:

- Several courses of antibiotics AND
- Trial of inhaled and/or oral steroids AND
- Allergy assessment and treatment when indicated

AND

- One or more of the following:
- Findings of obstruction of active infection on CT scan
- Symptomatic mucocele
- Negative CT scan but significant disease found on nasal endoscopy

OR

- C) Nasal polyposis causing or contributing to sinusitis

OR

- D) Complications of sinusitis including subperiosteal or orbital abscess, Pott's puffy tumor, brain abscess or meningitis

OR

- E) Invasive or allergic fungal sinusitis

OR

- F) Tumor of nasal cavity or sinuses

OR

- G) CSF rhinorrhea

[Adenoidectomy \(CPT 42830, 428305\) is included on line 470 only for treatment of children with chronic sinusitis who fail appropriate medical therapy.](#)

GUIDELINE NOTE 118, OBSTRUCTIVE SLEEP APNEA DIAGNOSIS AND TREATMENT FOR CHILDREN

Line 206

Obstructive sleep apnea (OSA) in children (18 or younger) must be diagnosed by

1. nocturnal polysomnography with an AHI >5 episodes/h or AHI>1 episodes/h with history and exam consistent with OSA, OR
2. nocturnal pulse oximetry with 3 or more SpO2 drops <90% and 3 or more clusters of desaturation events, or alternatives desaturation (>3%) index >3.5 episodes/h, OR

Adenoidectomy

3. use of a validated questionnaire (such as the Pediatric Sleep Questionnaire or OSA 18),
OR
4. consultation with a sleep medicine specialist.

Polysomnography and/or consultation with a sleep medicine specialist to support the diagnosis of OSA and/or to identify perioperative risk is recommended for

1. high risk children (i.e. children with cranio-facial abnormalities, neuromuscular disorders, Down syndrome, etc.)
2. children with equivocal indications for adenotonsillectomy (such as discordance between tonsillar size on physical examination and the reported severity of sleep-disordered breathing),
3. children younger than three years of age

Adenotonsillectomy is an appropriate first line treatment for children with OSA. Weight loss is recommended in addition to other therapy in patients who are overweight or obese.

Adenoidectomy without tonsillectomy is only covered when a child with OSA has previously had a tonsillectomy, when tonsillectomy is contraindicated, or when tonsillar hypertrophy is not present. More complex surgical treatments are only included on this line for children with craniofacial anomalies.

Intranasal corticosteroids are an option for children with mild OSA in whom adenotonsillectomy is contraindicated or for mild postoperative OSA.

CPAP is covered for a 3 month trial for children through age 18 who have

1. undergone surgery or are not candidates for surgery, AND
2. have documented residual sleep apnea symptoms (sleep disruption and/or significant desaturations) with residual daytime symptoms (daytime sleepiness or behavior problems)

CPAP will be covered for children through age 18 on an ongoing basis if:

- There is documentation of improvement in sleep disruption and daytime sleepiness and behavior problems with CPAP use
- Annual re-evaluation for CPAP demonstrates ongoing clinical benefit and compliance with use, defined as use of CPAP for at least four hours per night on 70% of the nights in a consecutive 30 day period

Adenoidectomy for recurrent or chronic nasal symptoms in children (Review)

van den Aardweg MTA, Schilder AGM, Herkert E, Boonacker CWB, Rovers MM



**THE COCHRANE
COLLABORATION®**

This is a reprint of a Cochrane review, prepared and maintained by The Cochrane Collaboration and published in *The Cochrane Library* 2010, Issue 3

<http://www.thecochranelibrary.com>



TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
PLAIN LANGUAGE SUMMARY	2
BACKGROUND	2
OBJECTIVES	3
METHODS	3
RESULTS	4
Figure 1.	6
Figure 2.	6
DISCUSSION	7
AUTHORS' CONCLUSIONS	7
ACKNOWLEDGEMENTS	8
REFERENCES	8
CHARACTERISTICS OF STUDIES	10
DATA AND ANALYSES	16
APPENDICES	16
WHAT'S NEW	17
HISTORY	18
CONTRIBUTIONS OF AUTHORS	18
DECLARATIONS OF INTEREST	18
SOURCES OF SUPPORT	18
DIFFERENCES BETWEEN PROTOCOL AND REVIEW	19
INDEX TERMS	19

[Intervention Review]

Adenoidectomy for recurrent or chronic nasal symptoms in children

Maaïke TA van den Aardweg¹, Anne GM Schilder², Ellen Herkert¹, Chantal WB Boonacker³, Maroeska M Rovers²

¹Department of Otorhinolaryngology, University Medical Center Utrecht, Utrecht, Netherlands. ²Department of Otorhinolaryngology & Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands. ³Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, Netherlands

Contact address: Maaïke TA van den Aardweg, Department of Otorhinolaryngology, University Medical Center Utrecht, Wilhelmina Children's Hospital, HP: KE.04.140.5, PO Box 85090, Utrecht, 3508 AB, Netherlands. M.T.A.vandenAardweg@umcutrecht.nl.

Editorial group: Cochrane Ear, Nose and Throat Disorders Group.

Publication status and date: Edited (no change to conclusions), published in Issue 3, 2010.

Review content assessed as up-to-date: 29 March 2009.

Citation: van den Aardweg MTA, Schilder AGM, Herkert E, Boonacker CWB, Rovers MM. Adenoidectomy for recurrent or chronic nasal symptoms in children. *Cochrane Database of Systematic Reviews* 2010, Issue 1. Art. No.: CD008282. DOI: 10.1002/14651858.CD008282.

Copyright © 2010 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

Background

Adenoidectomy, surgical removal of the adenoids, is a common ENT operation worldwide in children with recurrent or chronic nasal symptoms. A systematic review on the effectiveness of adenoidectomy in this specific group has not previously been performed.

Objectives

To assess the effectiveness of adenoidectomy versus non-surgical management in children with recurrent or chronic nasal symptoms.

Search strategy

We searched the Cochrane Ear, Nose and Throat Disorders Group Trials Register; the Cochrane Central Register of Controlled Trials (CENTRAL); PubMed; EMBASE; CINAHL; Web of Science; BIOSIS Previews; Cambridge Scientific Abstracts; mRCT and additional sources for published and unpublished trials. The date of the most recent search was 30 March 2009.

Selection criteria

Randomised controlled trials comparing adenoidectomy, with or without tympanostomy tubes, versus non-surgical management or tympanostomy tubes alone in children with recurrent or chronic nasal symptoms. The primary outcome studied was the number of episodes, days per episode and per year with nasal symptoms and the proportion of children with recurrent episodes of nasal symptoms. Secondary outcomes were mean number of episodes, mean number of days per episode and per year, and proportion of children with nasal obstruction alone.

Data collection and analysis

Two authors assessed trial quality and extracted data independently.

Main results

Only one study included children scheduled for adenoidectomy because of recurrent or chronic nasal symptoms or middle ear disease. In this study no beneficial effect of adenoidectomy was found. The numbers in this study were, however, small ($n = 76$) and the quality of the study was moderate. The outcome was improvement in episodes of common colds. The risk differences were non-significant, being 2% (95% CI -18% to 22%) and -11% (95% CI -28% to 7%) after 12 and 24 months, respectively.

A second study included children with recurrent acute otitis media ($n = 180$). As otitis media is known to be associated with nasal symptoms, the number of days with rhinitis was studied as a secondary outcome measure. The risk difference was non-significant, being -4 days (95% CI -13 to 7 days).

Authors' conclusions

Current evidence regarding the effect of adenoidectomy on recurrent or chronic nasal symptoms or nasal obstruction alone is sparse, inconclusive and has a significant risk of bias.

High quality trials assessing the effectiveness of adenoidectomy in children with recurrent or chronic nasal symptoms should be initiated.

PLAIN LANGUAGE SUMMARY

Adenoidectomy for recurrent of chronic nasal symptoms in children

Infections of the upper respiratory tract, presenting as recurrent nasal symptoms (nasal discharge with or without nasal obstruction) are very common in children. Removal of the adenoids (adenoidectomy) is a surgical procedure that is frequently performed in these children. It is thought that adenoidectomy prevents recurrence of nasal symptoms.

Our review, which includes two studies (256 children), shows that it is uncertain whether adenoidectomy is effective in children with recurrent or chronic nasal symptoms. Further high quality trials are needed.

BACKGROUND

Incidence

Adenoidectomy is one of the most frequently performed surgical procedures in children in Western countries. Annual adenoidectomy rates, however, differ between countries, from 127/10,000 children per year in Belgium, 101/10,000 children per year in the Netherlands and 39/10,000 children per year in England to 24/10,000 and 17/10,000 children per year in the United States and Canada, respectively (Schilder 2004).

Adenoidectomy

Indications for adenoidectomy include recurrent or chronic nasal discharge, recurrent episodes of acute otitis media (AOM), persistent otitis media with effusion (OME) and symptoms of upper airway obstruction. In most cases children are operated on for a combination of nasal and middle ear symptoms. The operation

involves removing the adenoids - a nasopharyngeal reservoir of potential respiratory pathogens and a potential cause of obstruction of the nasal airway. As such, nasal breathing is thought to improve and recurrence of nasal discharge is thought to decrease.

Evidence for adenoidectomy

At present the effectiveness of adenoidectomy in children with recurrent or chronic nasal symptoms remains uncertain and practice is experience-based rather than evidence-based. Some surgeons prefer to perform adenoidectomy in these children, whereas others do not. In previous systematic reviews for *The Cochrane Library* the effectiveness of 1) grommets (ventilation tubes) for recurrent acute otitis media in children (McDonald 2008), 2) grommets (ventilation tubes) for hearing loss associated with otitis media with effusion in children (Lous 2005) and 3) tonsillectomy or adeno-tonsillectomy versus non-surgical treatment for chronic/recurrent acute tonsillitis (Burton 2009) have all been assessed. For adenoidectomy in children with nasal symptoms no such review



Adenoidectomy outcomes in pediatric rhinosinusitis: A meta-analysis^{☆,☆☆}

Scott E. Brietzke^{a,*}, Matthew T. Brigger^b

^a Department of Otolaryngology, Walter Reed Army Medical Center, 6900 Georgia Avenue, Washington, DC 20307, United States

^b Department of Otolaryngology, National Naval Medical Center, Bethesda, MD, United States

Received 11 June 2008; received in revised form 12 July 2008; accepted 14 July 2008

KEYWORDS

Adenoidectomy;
Pediatric rhinosinusitis;
Otolaryngology;
Evidence-based
medicine

Summary

Objective: To evaluate the currently available literature regarding the reported effectiveness of adenoidectomy alone in the management of medically refractory pediatric chronic rhinosinusitis.

Methods: The MEDLINE and EMBASE databases were systematically searched for English language manuscripts reporting on effectiveness of adenoidectomy alone in management of medically refractory pediatric rhinosinusitis. Additional manuscripts were identified by manual searching. Random effects modeling was performed to produce summary estimates of adenoidectomy effectiveness.

Results: Nine studies met the inclusion criteria. Six were cohort studies (level 2b) and four were case series (level 4). Mean sample size was 46 subjects (range = 10–121) with grand mean age of 5.8 years (range 4.4–6.9 years). All studies showed that sinusitis symptoms or outcomes improved in half or more patients after adenoidectomy. Eight studies were sufficiently similar to undergo meta-analysis. The summary estimate of the proportion of patients who significantly improved after adenoidectomy was 69.3% (95% CI = 56.8–81.7%, $p < 0.001$). The possibility of author bias was explored as one author group contributed a large proportion of patients to the study group.

Conclusion: Adenoidectomy reduces caregiver reported symptoms of chronic rhinosinusitis in a majority of pediatric patients. Given its simplicity, low risk profile, and apparent effectiveness, adenoidectomy should be considered first line therapy for medically refractory, uncomplicated pediatric rhinosinusitis.

Published by Elsevier Ireland Ltd.

[☆] The views herein are the private views of the authors and do not reflect the views of the Department of the Army, the Department of the Navy or the Department of Defense.

^{☆☆} Presented at the American Society of Pediatric Otolaryngology Annual Meeting San Diego, CA, USA, May 2007.

* Corresponding author. Fax: +1 202 782 4647.

E-mail address: SEBrietzke@msn.com (S.E. Brietzke).

The effect of adenotonsillectomy in children with OSA

Tomoko Shintani *, Kohji Asakura, Akikatsu Kataura

Department of Otolaryngology, Sapporo Medical University, S1 W16 Chuo-ku Sapporo, 060 Japan

Received 15 July 1997; received in revised form 3 December 1997; accepted 15 March 1998

Abstract

Adenotonsillar hypertrophy and abnormal facial morphology are thought to be important for the occurrence of obstructive sleep apnea syndrome (OSA). We evaluated the effects of adenoidectomy and/or tonsillectomy and the relationship between the treatment results and facial morphology in 134 children with OSA. Significant improvements in apnea-hypopnea index (AHI) and lowest blood oxygen saturation (SaO_2) were noted and 78.5% of the patients improved after adenoidectomy and/or tonsillectomy. Additional operations were needed in two out of 13 cases of the adenoidectomy group and two out of four cases of the adeno-monotonsillectomy group. In the adenotonsillectomy group, the unimproved children tended to have smaller tonsils, narrower epipharyngeal airspace, and more poorly-developed maxillary and mandibular protrusion than the improved children. © 1998 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Obstructive sleep apnea; Adenoidectomy; Tonsillectomy; Cephalometric analysis

1. Introduction

Obstructive sleep apnea (OSA) has been recognized as an important entity in children [3]. The symptoms of OSA are snoring, apnea, and open mouth [5]. Continuity of upper airway obstruction may lead to abnormal facial morphology [13], chest deformation [7] and growth retardation [5,11,15]. In severe cases, arrhythmia, right-sided heart failure, and cor pulmonale (pulmonary hy-

pertension) may also be noted [1,8,9,11,16], and in some cases, they can cause death. Adenotonsillar hypertrophy is the most common cause of obstructive sleep apnea in children [9,16]. In children with OSA, adenoidectomy and/or tonsillectomy are often the initial treatment [2]. The symptoms of OSA such as snoring, apnea, mouth breathing and poor school performance have been shown by several studies to be improved following adenotonsillectomy [9,17]. However, not every child with large tonsils and adenoids has symptoms of OSA. In our previous study [14], it was suspected that the abnormal facial morphology in OSA

* Corresponding author. Tel: +81 11 6112111 (3491); fax: +81 11 6155405.

JOURNAL CLUB

Edited by Patrina Caldwell (patrinac@chw.edu.au)

How effective is adenoidectomy alone for treatment of obstructive sleep apnoea in a child who presents with adenoid hypertrophy?

Rebecca Hensley, Rita Machaalani, Alan Cheng and Karen Waters

The Children's Hospital at Westmead, Westmead and The University of Sydney, Sydney, New South Wales, Australia



Clinical Scenario

A girl aged 20 months presents with parental concern about loud snoring and increased work of breathing when sleeping. On examination, she has adenoid facies (a long, narrow face with mouth open¹ and small Grade 1/4 tonsils (using the Brodsky Grading System)).² A lateral neck X-ray shows enlarged adenoids obstructing the nasal airway. The child has a history of two ear infections, both of which occurred in the 12 months prior to presentation, and both are resolved after a single course of oral antibiotics. There is no family or personal history of atopy, and skin prick testing was negative for common allergens. Nasal steroid spray had been previously prescribed, but after her parents found this, which led to little improvement, they stopped the treatment. An overnight sleep study showed mild obstructive sleep apnoea (OSA) with an index of obstructive/mixed respiratory events of 9.2 events/h, a minimum oxygen saturation of 87% and no carbon dioxide retention.

Our ear, nose and throat (ENT) surgeon reviewed the 20-month girl and suggested that the most appropriate surgery would be adenoidectomy without tonsillectomy.

Structured clinical question

'In children with OSA and adenoid hypertrophy without tonsil hypertrophy (P) does adenoidectomy alone (I) comparing sleep studies (polysomnography (PSG)) before and after surgery (C) provide acute relief of OSA (O)'.

Search Strategy

The OVID database was searched from the years 1950 to the present. The Medical Subject Headings' terms 'adenoidectomy' and 'pediatric' and 'airway obstruction/or sleep disorders/or sleep apnea, obstructive/or sleep apnea syndromes/or sleep/or polysomnography' were combined. Key journals (such as *SLEEP*) and conference proceedings (Australasian Sleep Association 2006–2008) were hand searched, and pertinent grey data were included in the search. The search was conducted in this manner recommended by evidence-based practice guidelines.

Correspondence: Prof Karen Waters, Respiratory Support Service, The Children's Hospital at Westmead, 212 Hawkesbury Road, Westmead, NSW 2145, Australia. Fax: +61 29845 3396; email: karenw2@chw.edu.au

Accepted for publication 12 June 2011.

The search resulted in 226 articles, which were reviewed by title and/or abstract, to ascertain if the study investigated adenoidectomy alone, as a separate procedure to adenotonsillectomy, as the primary intervention. Only publications in English were reviewed. A final 38 papers addressed adenoidectomy as a separate procedure; 25 were relevant to the clinical question, and of these, the seven studies that included evaluation of adenoidectomy alone were most appropriate and are tabulated later. This technique facilitated a rigorous, unbiased review of the literature (Table 1). The last three studies in the table examined longer-term requirements for additional surgery after adenoidectomy.^{10–12}

What is the best paper to answer my question?

The study by Shintani et al. in 1998⁸ was chosen because it included the most comprehensive analysis of obstructive apnoea in a group of patients with studies undertaken before and after adenoidectomy. The study included 134 ENT surgery candidates, aged 1–9 years of age, and provided specific data about 13 children who had adenoidectomy alone. Although the primary outcome variables were cephalometrics to measure facial dimensions, the study evaluated the short-term outcome of adenoidectomy on OSA with 2 months follow-up after the procedure. The severity of the sleep apnoea was quantified for children undergoing adenoidectomy alone using a modified PSG that included respiratory inductive plethysmography (trace of chest wall movement with respiratory efforts) and SaO₂, performed pre-operatively and then 2 months post-operatively. Children who underwent adenoidectomy alone showed significant reduction in apnoea hypopnoea index (AHI) and in SaO₂ nadir. The study found significant ($P < 0.005$) improvement in AHI and SaO₂ after adenoidectomy alone. The rate of improvement in AHI and lowest SaO₂ was higher for children who underwent adenoidectomy (84.5%) than for those who underwent adenotonsillectomy (75.4%). Facial morphology was not analysed in the children who underwent adenoidectomy alone. ENT surgery was required for two adenoidectomy patients, who were both 2 years of age.

Using the Critical Appraisal Toolkit for Cohort Studies,¹³ we found that the study asked a clearly focused question that was applicable to the children undergoing adenoidectomy alone, as a subgroup. A cohort study is a good way of answering the question. Only children with a pre-operative PSG result demonstrating an AHI > 10/h were recruited into the OSA arm of the study. Exclusion criteria included developmental delay,

PEDIATRICS®

OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

Diagnosis and Management of Childhood Obstructive Sleep Apnea Syndrome

Carole L. Marcus, Lee Jay Brooks, Kari A. Draper, David Gozal, Ann Carol Halbower, Jacqueline Jones, Michael S. Schechter, Stephen Howard Sheldon, Karen Spruyt, Sally Davidson Ward, Christopher Lehmann and Richard N. Shiffman

Pediatrics; originally published online August 27, 2012;

DOI: 10.1542/peds.2012-1671

The online version of this article, along with updated information and services, is located on the World Wide Web at:

<http://pediatrics.aappublications.org/content/early/2012/08/22/peds.2012-1671>

PEDIATRICS is the official journal of the American Academy of Pediatrics. A monthly publication, it has been published continuously since 1948. PEDIATRICS is owned, published, and trademarked by the American Academy of Pediatrics, 141 Northwest Point Boulevard, Elk Grove Village, Illinois, 60007. Copyright © 2012 by the American Academy of Pediatrics. All rights reserved. Print ISSN: 0031-4005. Online ISSN: 1098-4275.

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN™





CLINICAL PRACTICE GUIDELINE

Diagnosis and Management of Childhood Obstructive Sleep Apnea Syndrome

abstract

FREE

OBJECTIVES: This revised clinical practice guideline, intended for use by primary care clinicians, provides recommendations for the diagnosis and management of the obstructive sleep apnea syndrome (OSAS) in children and adolescents. This practice guideline focuses on uncomplicated childhood OSAS, that is, OSAS associated with adenotonsillar hypertrophy and/or obesity in an otherwise healthy child who is being treated in the primary care setting.

METHODS: Of 3166 articles from 1999–2010, 350 provided relevant data. Most articles were level II–IV. The resulting evidence report was used to formulate recommendations.

RESULTS AND CONCLUSIONS: The following recommendations are made. (1) All children/adolescents should be screened for snoring. (2) Polysomnography should be performed in children/adolescents with snoring and symptoms/signs of OSAS; if polysomnography is not available, then alternative diagnostic tests or referral to a specialist for more extensive evaluation may be considered. (3) Adenotonsillectomy is recommended as the first-line treatment of patients with adenotonsillar hypertrophy. (4) High-risk patients should be monitored as inpatients postoperatively. (5) Patients should be reevaluated postoperatively to determine whether further treatment is required. Objective testing should be performed in patients who are high risk or have persistent symptoms/signs of OSAS after therapy. (6) Continuous positive airway pressure is recommended as treatment if adenotonsillectomy is not performed or if OSAS persists postoperatively. (7) Weight loss is recommended in addition to other therapy in patients who are overweight or obese. (8) Intranasal corticosteroids are an option for children with mild OSAS in whom adenotonsillectomy is contraindicated or for mild postoperative OSAS. *Pediatrics* 2012;130:576–584

INTRODUCTION

Obstructive sleep apnea syndrome (OSAS) is a common condition in childhood and can result in severe complications if left untreated. In 2002, the American Academy of Pediatrics (AAP) published a practice guideline for the diagnosis and management of childhood OSAS.¹ Since that time, there has been a considerable increase in publications and research on the topic; thus, the guidelines have been revised.

Carole L. Marcus, MBBCh, Lee Jay Brooks, MD, Kari A. Draper, MD, David Gozal, MD, Ann Carol Halbower, MD, Jacqueline Jones, MD, Michael S. Schechter, MD, MPH, Stephen Howard Sheldon, DO, Karen Spruyt, PhD, Sally Davidson Ward, MD, Christopher Lehmann, MD, Richard N. Shiffman, MD

KEY WORDS

snoring, sleep-disordered breathing, adenotonsillectomy, continuous positive airway pressure

ABBREVIATIONS

AAP—American Academy of Pediatrics

AHI—apnea hypopnea index

CPAP—continuous positive airway pressure

OSAS—obstructive sleep apnea syndrome

This document is copyrighted and is property of the American Academy of Pediatrics and its Board of Directors. All authors have filed conflict of interest statements with the American Academy of Pediatrics. Any conflicts have been resolved through a process approved by the Board of Directors. The American Academy of Pediatrics has neither solicited nor accepted any commercial involvement in the development of the content of this publication.

The recommendations in this report do not indicate an exclusive course of treatment or serve as a standard of medical care. Variations, taking into account individual circumstances, may be appropriate.

All clinical practice guidelines from the American Academy of Pediatrics automatically expire 5 years after publication unless reaffirmed, revised, or retired at or before that time.

www.pediatrics.org/cgi/doi/10.1542/peds.2012-1671

doi:10.1542/peds.2012-1671

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

Copyright © 2012 by the American Academy of Pediatrics

The purposes of this revised clinical practice guideline are to (1) increase the recognition of OSAS by primary care clinicians to minimize delay in diagnosis and avoid serious sequelae of OSAS; (2) evaluate diagnostic techniques; (3) describe treatment options; (4) provide guidelines for follow-up; and (5) discuss areas requiring further research. The recommendations in this statement do not indicate an exclusive course of treatment. Variations, taking into account individual circumstances, may be appropriate.

This practice guideline focuses on uncomplicated childhood OSAS—that is, the OSAS associated with adenotonsillar hypertrophy and/or obesity in an otherwise healthy child who is being treated in the primary care setting. This guideline specifically excludes infants younger than 1 year of age, patients with central apnea or hypoventilation syndromes, and patients with OSAS associated with other medical disorders, including but not limited to Down syndrome, craniofacial anomalies, neuromuscular disease (including cerebral palsy), chronic lung disease, sickle cell disease, metabolic disease, or laryngomalacia. These important patient populations are too complex to discuss within the scope of this article and require consultation with a pediatric subspecialist.

Additional information providing justification for the key action statements and a detailed review of the literature are provided in the accompanying technical report available online.²

METHODS OF GUIDELINE DEVELOPMENT

Details of the methods of guideline development are included in the accompanying technical report.² The AAP selected a subcommittee composed of pediatricians and other experts in the fields of sleep medicine, pulmonology, and otolaryngology, as well as experts

from epidemiology and pediatric practice to develop an evidence base of literature on this topic. The committee included liaison members from the AAP Section on Otolaryngology-Head and Neck Surgery, American Thoracic Society, American Academy of Sleep Medicine, American College of Chest Physicians, and the National Sleep Foundation. Committee members signed forms disclosing conflicts of interest.

An automated search of the literature on childhood OSAS from 1999 to 2008 was performed by using 5 scientific literature search engines.² The medical subject heading terms that were used in all fields were snoring, apnea, sleep-disordered breathing, sleep-related breathing disorders, upper airway resistance, polysomnography, sleep study, adenoidectomy, tonsillectomy, continuous positive airway pressure, obesity, adiposity, hypopnea, hypoventilation, cognition, behavior, and neuropsychology. Reviews, case reports, letters to the editor, and abstracts were not included. Non-English-language articles, animal studies, and studies relating to infants younger than 1 year and to special populations (eg, children with craniofacial anomalies or sickle cell disease) were excluded. In several steps, a total of 3166 hits was reduced to 350 articles, which underwent detailed review.² Committee members selectively updated this literature search for articles published from 2008 to 2011 specific to guideline categories. Details of the literature grading system are available in the accompanying technical report.

Since publication of the previous guidelines, there has been an improvement in the quality of OSAS studies in the literature; however, there remain few randomized, blinded, controlled studies. Most studies were questionnaire or polysomnography based. Many studies used standard definitions for pediatric polysomnography scoring, but

the interpretation of polysomnography (eg, the apnea hypopnea index [AHI] criterion used for diagnosis or to determine treatment) varied widely. The guideline notes the quality of evidence for each key action statement. Additional details are available in the technical report.

The evidence-based approach to guideline development requires that the evidence in support of each key action statement be identified, appraised, and summarized and that an explicit link between evidence and recommendations be defined. Evidence-based recommendations reflect the quality of evidence and the balance of benefit and harm that is anticipated when the recommendation is followed. The AAP policy statement, “Classifying Recommendations for Clinical Practice Guidelines,”³ was followed in designating levels of recommendation (see Fig 1 and Table 1).

DEFINITION

This guideline defines OSAS in children as a “disorder of breathing during sleep characterized by prolonged partial upper airway obstruction and/or intermittent complete obstruction (obstructive apnea) that disrupts normal ventilation during sleep and normal sleep patterns,”⁴ accompanied by symptoms or signs, as listed in Table 2. Prevalence rates based on level I and II studies range from 1.2% to 5.7%.^{5–7} Symptoms include habitual snoring (often with intermittent pauses, snorts, or gasps), disturbed sleep, and daytime neurobehavioral problems. Daytime sleepiness may occur, but is uncommon in young children. OSAS is associated with neurocognitive impairment, behavioral problems, failure to thrive, hypertension, cardiac dysfunction, and systemic inflammation. Risk factors include adenotonsillar hypertrophy, obesity, craniofacial anomalies, and neuromuscular disorders. Only the first 2 risk factors are

Evidence Quality	Preponderance of Benefit or Harm	Balance of Benefit and Harm
A. Well designed RCTs or diagnostic studies on relevant population	Strong Recommendation	Option
B. RCTs or diagnostic studies with minor limitations;overwhelmingly consistent evidence from observational studies	Recommendation	
C. Observational studies (case-control and cohort design)	Option	
D. Expert opinion, case reports, reasoning from first principles	Option	No Rec
X. Exceptional situations where validating studies cannot be performed and there is a clear preponderance of benefit or harm	Strong Recommendation	
	Recommendation	

FIGURE 1

Evidence quality. Integrating evidence quality appraisal with an assessment of the anticipated balance between benefits and harms if a policy is carried out leads to designation of a policy as a strong recommendation, recommendation, option, or no recommendation. RCT, randomized controlled trial; Rec, recommendation.

discussed in this guideline. In this guideline, obesity is defined as a BMI >95th percentile for age and gender.⁸

KEY ACTION STATEMENTS

Key Action Statement 1: Screening for OSAS

As part of routine health maintenance visits, clinicians should inquire whether the child or adolescent snores. If the answer is affirmative or if a child or adolescent presents with signs or symptoms of OSAS (Table 2), clinicians should perform

a more focused evaluation. (Evidence Quality: Grade B, Recommendation Strength: Recommendation.)

Evidence Profile KAS 1

- Aggregate evidence quality: B
- Benefit: Early identification of OSAS is desirable, because it is a high-prevalence condition, and identification and treatment can result in alleviation of current symptoms, improved quality of life, prevention of sequelae, education of parents, and decreased health care utilization.

- Harm: Provider time, patient and parent time.
- Benefits-harms assessment: Preponderance of benefit over harm.
- Value judgments: Panelists believe that identification of a serious medical condition outweighs the time expenditure necessary for screening.
- Role of patient preferences: None.
- Exclusions: None.
- Intentional vagueness: None.
- Strength: Recommendation.

Almost all children with OSAS snore,^{9–11} although caregivers frequently do not volunteer this information at medical visits.¹² Thus, asking about snoring at each health maintenance visit (as well as at other appropriate times, such as when evaluating for tonsillitis) is a sensitive, albeit nonspecific, screening measure that is quick and easy to perform. Snoring is common in children and adolescents; however, OSAS is less common. Therefore, an affirmative answer should be followed by a detailed history and examination to determine whether further evaluation for OSAS is needed (Table 2); this clinical evaluation alone

TABLE 1 Definitions and Recommendation Implications

Statement	Definition	Implication
Strong recommendation	A strong recommendation in favor of a particular action is made when the anticipated benefits of the recommended intervention clearly exceed the harms (as a strong recommendation against an action is made when the anticipated harms clearly exceed the benefits) and the quality of the supporting evidence is excellent. In some clearly identified circumstances, strong recommendations may be made when high-quality evidence is impossible to obtain and the anticipated benefits strongly outweigh the harms.	Clinicians should follow a strong recommendation unless a clear and compelling rationale for an alternative approach is present.
Recommendation	A recommendation in favor of a particular action is made when the anticipated benefits exceed the harms but the quality of evidence is not as strong. Again, in some clearly identified circumstances, recommendations may be made when high-quality evidence is impossible to obtain but the anticipated benefits outweigh the harms.	It would be prudent for clinicians to follow a recommendation, but they should remain alert to new information and sensitive to patient preferences.
Option	Options define courses that may be taken when either the quality of evidence is suspect or carefully performed studies have shown little clear advantage to one approach over another.	Clinicians should consider the option in their decision-making, and patient preference may have a substantial role.
No recommendation	No recommendation indicates that there is a lack of pertinent published evidence and that the anticipated balance of benefits and harms is presently unclear.	Clinicians should be alert to new published evidence that clarifies the balance of benefit versus harm.

TABLE 2 Symptoms and Signs of OSAS

History
Frequent snoring (≥ 3 nights/wk)
Labored breathing during sleep
Gasps/snorting noises/observed episodes of apnea
Sleep enuresis (especially secondary enuresis) ^a
Sleeping in a seated position or with the neck hyperextended
Cyanosis
Headaches on awakening
Daytime sleepiness
Attention-deficit/hyperactivity disorder
Learning problems
Physical examination
Underweight or overweight
Tonsillar hypertrophy
Adenoidal facies
Micrognathia/retrognathia
High-arched palate
Failure to thrive
Hypertension

^a Enuresis after at least 6 mo of continence.

does not establish the diagnosis (see technical report). Occasional snoring, for example, with an upper respiratory tract infection, is less of a concern than snoring that occurs at least 3 times a week and is associated with any of the symptoms or signs listed in Table 2.

Key Action Statement 2A: Polysomnography

If a child or adolescent snores on a regular basis and has any of the complaints or findings shown in Table 2, clinicians should either (1) obtain a polysomnogram (Evidence Quality A, Key Action strength: Recommendation) OR (2) refer the patient to a sleep specialist or otolaryngologist for a more extensive evaluation (Evidence quality D, Key Action strength: Option). (Evidence Quality: Grade A for polysomnography; Grade D for specialist referral, Recommendation Strength: Recommendation.)

Evidence Profile KAS 2A: Polysomnography

- Aggregate evidence quality: A
- Benefits: Establish diagnosis and determine severity of OSAS.

- Harm: Expense, time, anxiety/discomfort.
- Benefits-harms assessment: Preponderance of benefit over harm.
- Value judgments: Panelists weighed the value of establishing a diagnosis as more important than the minor potential harms listed.
- Role of patient preferences: Small because of preponderance of evidence that polysomnography is the most accurate way to make a diagnosis.
- Exclusions: See Key Action Statement 2B regarding lack of availability.
- Intentional vagueness: None.
- Strength: Recommendation.

Evidence Profile KAS 2A: Referral

- Aggregate evidence quality: D
- Benefits: Subspecialist may be better able to establish diagnosis and determine severity of OSAS.
- Harm: Expense, time, anxiety/discomfort.
- Benefits-harms assessment: Preponderance of benefit over harm.
- Value judgments: Panelists weighed the value of establishing a diagnosis as more important than the minor potential harms listed.
- Role of patient preferences: Large.
- Exclusions: None.
- Intentional vagueness: None.
- Strength: Option.

Although history and physical examination are useful to screen patients and determine which patients need further investigation for OSAS, the sensitivity and specificity of the history and physical examination are poor (see accompanying technical report). Physical examination when the child is awake may be normal, and the size of the tonsils cannot be used to predict the presence of OSAS in an individual child. Thus, objective testing is required. The gold standard test

is overnight, attended, in-laboratory polysomnography (sleep study). This is a noninvasive test involving the measurement of a number of physiologic functions overnight, typically including EEG; pulse oximetry; oronasal airflow, abdominal and chest wall movements, partial pressure of carbon dioxide (P_{CO_2}); and video recording.¹³ Specific pediatric measuring and scoring criteria should be used.¹³ Polysomnography will demonstrate the presence or absence of OSAS. Polysomnography also demonstrates the severity of OSAS, which is helpful in planning treatment and in postoperative short- and long-term management.

Key Action Statement 2B: Alternative Testing

If polysomnography is not available, then clinicians may order alternative diagnostic tests, such as nocturnal video recording, nocturnal oximetry, daytime nap polysomnography, or ambulatory polysomnography. (Evidence Quality: Grade C, Recommendation Strength: Option.)

Evidence Profile KAS 2B

- Aggregate evidence quality: C
- Benefit: Varying positive and negative predictive values for establishing diagnosis.
- Harm: False-negative and false-positive results may underestimate or overestimate severity, expense, time, anxiety/discomfort.
- Benefits-harms assessment: Equilibrium of benefits and harms.
- Value judgments: Opinion of the panel that some objective testing is better than none. Pragmatic decision based on current shortage of pediatric polysomnography facilities (this may change over time).
- Role of patient preferences: Small, if choices are limited by availability;

families may choose to travel to centers where more extensive facilities are available.

- Exclusions: None.
- Intentional vagueness: None.
- Strength: Option.

Although polysomnography is the gold standard for diagnosis of OSAS, there is a shortage of sleep laboratories with pediatric expertise. Hence, polysomnography may not be readily available in certain regions of the country. Alternative diagnostic tests have been shown to have weaker positive and negative predictive values than polysomnography, but nevertheless, objective testing is preferable to clinical evaluation alone. If an alternative test fails to demonstrate OSAS in a patient with a high pretest probability, full polysomnography should be sought.

Key Action Statement 3: Adenotonsillectomy

If a child is determined to have OSAS, has a clinical examination consistent with adenotonsillar hypertrophy, and does not have a contraindication to surgery (see Table 3), the clinician should recommend adenotonsillectomy as the first line of treatment. If the child has OSAS but does not have adenotonsillar hypertrophy, other treatment should be considered (see Key Action Statement 6). Clinical judgment is required to determine the benefits of adenotonsillectomy compared with other treatments in obese children with varying degrees of adenotonsillar hypertrophy. (Evidence Quality: Grade B, Recommendation Strength: Recommendation.)

Evidence Profile KAS 3

- Aggregate evidence quality: B
- Benefit: Improve OSAS and accompanying symptoms and sequelae.

- Harm: Pain, anxiety, dehydration, anesthetic complications, hemorrhage, infection, postoperative respiratory difficulties, velopharyngeal incompetence, nasopharyngeal stenosis, death.
- Benefits-harms assessment: Preponderance of benefit over harm.
- Value judgments: The panel sees the benefits of treating OSAS as more beneficial than the low risk of serious consequences.
- Role of patient preferences: Low; continuous positive airway pressure (CPAP) is an option but involves prolonged, long-term treatment as compared with a single, relatively low-risk surgical procedure.
- Exclusions: See Table 3.
- Intentional vagueness: None.
- Strength: Recommendation.

Adenotonsillectomy is very effective in treating OSAS. Adenoidectomy or tonsillectomy alone may not be sufficient, because residual lymphoid tissue may contribute to persistent obstruction. In otherwise healthy children with adenotonsillar hypertrophy, adenotonsillectomy is associated with improvements in symptoms and sequelae of OSAS. Postoperative polysomnography typically shows a major decrease in the number of obstructive events, although some obstructions may still be present. Although obese children may have less satisfactory results, many will be adequately treated with

TABLE 3 Contraindications for Adenotonsillectomy

Absolute contraindications
No adenotonsillar tissue (tissue has been surgically removed)
Relative contraindications
Very small tonsils/adenoid
Morbid obesity and small tonsils/adenoid
Bleeding disorder refractory to treatment
Submucous cleft palate
Other medical conditions making patient medically unstable for surgery

adenotonsillectomy; however, further research is needed to determine which obese children are most likely to benefit from surgery. In this population, the benefits of a 1-time surgical procedure, with a small but real risk of complications, need to be weighed against long-term treatment with CPAP, which is associated with discomfort, disruption of family lifestyle, and risks of poor adherence. Potential complications of adenotonsillectomy are shown in Table 4. Although serious complications (including death) may occur, the rate of these complications is low, and the risks of complications need to be weighed against the consequences of untreated OSAS. In general, a 1-time only procedure with a relatively low morbidity is preferable to lifelong treatment with CPAP; furthermore, the efficacy of CPAP is limited by generally suboptimal adherence. Other treatment options, such as anti-inflammatory medications, weight loss, or tracheostomy, are less effective, are difficult to achieve, or have higher morbidity, respectively.

Key Action Statement 4: High-Risk Patients Undergoing Adenotonsillectomy

Clinicians should monitor high-risk patients (Table 5) undergoing adenotonsillectomy as inpatients postoperatively. (Evidence Quality: Grade B, Recommendation Strength: Recommendation.)

TABLE 4 Risks of Adenotonsillectomy

Minor
Pain
Dehydration attributable to postoperative nausea/vomiting and poor oral intake
Major
Anesthetic complications
Acute upper airway obstruction during induction or emergence from anesthesia
Postoperative respiratory compromise
Hemorrhage
Velopharyngeal incompetence
Nasopharyngeal stenosis
Death

TABLE 5 Risk Factors for Postoperative Respiratory Complications in Children With OSAS Undergoing Adenotonsillectomy

Younger than 3 y of age
Severe OSAS on polysomnography ^a
Cardiac complications of OSAS
Failure to thrive
Obesity
Craniofacial anomalies ^b
Neuromuscular disorders ^b
Current respiratory infection

^a It is difficult to provide exact polysomnographic criteria for severity, because these criteria will vary depending on the age of the child; additional comorbidities, such as obesity, asthma, or cardiac complications of OSAS; and other polysomnographic criteria that have not been evaluated in the literature, such as the level of hypercapnia and the frequency of desaturation (as compared with lowest oxygen saturation). Nevertheless, on the basis of published studies (primarily Level III, see Technical Report), it is recommended that all patients with a lowest oxygen saturation <80% (either on preoperative polysomnography or during observation in the recovery room postoperatively) or an AHI ≥ 24 /h be observed as inpatients postoperatively as they are at increased risk for postoperative respiratory compromise. Additionally, on the basis of expert consensus, it is recommended that patients with significant hypercapnia on polysomnography (peak $P_{CO_2} \geq 60$ mm Hg) be admitted postoperatively. The committee noted that that most published studies were retrospective and not comprehensive, and therefore these recommendations may change if higher-level studies are published. Clinicians may decide to admit patients with less severe polysomnographic abnormalities based on a constellation of risk factors (age, comorbidities, and additional polysomnographic factors) for a particular individual.

^b Not discussed in these guidelines.

Evidence Profile KAS 4

- Aggregate evidence quality: B
- Benefit: Effectively manage severe respiratory compromise and avoid death.
- Harm: Expense, time, anxiety.
- Benefits-harms assessment: Preponderance of benefit over harm.
- Value judgments: The panel believes that early recognition of any serious adverse events is critically important.
- Role of patient preferences: Minimal; this is an important safety issue.
- Exclusions: None.
- Intentional vagueness: None.
- Strength: Recommendation.

Patients with OSAS may develop respiratory complications, such as worsening

of OSAS or pulmonary edema, in the immediate postoperative period. Death attributable to respiratory complications in the immediate postoperative period has been reported in patients with severe OSAS. Identified risk factors are shown in Table 5. High-risk patients should undergo surgery in a center capable of treating complex pediatric patients. They should be hospitalized overnight for close monitoring postoperatively. Children with an acute respiratory infection on the day of surgery, as documented by fever, cough, and/or wheezing, are at increased risk of postoperative complications and, therefore, should be rescheduled or monitored closely postoperatively. Clinicians should decide on an individual basis whether these patients should be rescheduled, taking into consideration the severity of OSAS in the particular patient and keeping in mind that many children with adenotonsillar hypertrophy have chronic rhinorrhea and nasal congestion, even in the absence of viral infections.

Key Action Statement 5: Reevaluation

Clinicians should clinically reassess all patients with OSAS for persisting signs and symptoms after therapy to determine whether further treatment is required. (Evidence Quality: Grade B, Recommendation Strength: Recommendation.)

Evidence Profile KAS 5A

- Aggregate evidence quality: B
- Benefit: Determine effects of treatment.
- Harm: Expense, time.
- Benefits-harms assessment: Preponderance of benefit over harm.
- Value judgments: Data show that a significant proportion of children continue to have abnormalities postoperatively; therefore, the panel deter-

mined that the benefits of follow-up outweigh the minor inconveniences.

- Role of patient preferences: Minimal; follow-up is good clinical practice.
- Exclusions: None.
- Intentional vagueness: None.
- Strength: Recommendation.

Clinicians should reassess OSAS-related symptoms and signs (Table 2) after 6 to 8 weeks of therapy to determine whether further evaluation and treatment are indicated. Objective data regarding the timing of the postoperative evaluation are not available. Most clinicians recommend reevaluation 6 to 8 weeks after treatment to allow for healing of the operative site and to allow time for upper airway, cardiac, and central nervous system recovery. Patients who remain symptomatic should undergo objective testing (see Key Action Statement 2) or be referred to a sleep specialist for further evaluation.

Key Action Statement 5B: Reevaluation of High-Risk Patients

Clinicians should reevaluate high-risk patients for persistent OSAS after adenotonsillectomy, including those who had a significantly abnormal baseline polysomnogram, have sequelae of OSAS, are obese, or remain symptomatic after treatment, with an objective test (see Key Action Statement 2) or refer such patients to a sleep specialist. (Evidence Quality: Grade B, Recommendation Strength: Recommendation.)

Evidence Profile KAS 5B

- Aggregate evidence quality: B
- Benefit: Determine effects of treatment.
- Harm: Expense, time, anxiety/discomfort.
- Benefits-harms assessment: Preponderance of benefit over harm.

- Value judgments: Given the panel's concerns about the consequences of OSAS and the frequency of postoperative persistence in high-risk groups, the panel believes that the follow-up costs are outweighed by benefits of recognition of persistent OSAS. A minority of panelists believed that all children with OSAS should have follow-up polysomnography because of the high prevalence of persistent postoperative abnormalities on polysomnography, but most panelists believed that persistent polysomnographic abnormalities in uncomplicated children with mild OSAS were usually mild in patients who were asymptomatic after surgery.
- Role of patient preferences: Minimal. Further evaluation is needed to determine the need for further treatment.
- Exclusions: None.
- Intentional vagueness: None.
- Strength: Recommendation.

Numerous studies have shown that a large proportion of children at high risk continue to have some degree of OSAS postoperatively^{10,13,14}; thus, objective evidence is required to determine whether further treatment is necessary.

Key Action Statement 6: CPAP

Clinicians should refer patients for CPAP management if symptoms/signs (Table 2) or objective evidence of OSAS persists after adenotonsillectomy or if adenotonsillectomy is not performed. (Evidence Quality: Grade B, Recommendation Strength: Recommendation.)

Evidence Profile KAS 6

- Aggregate evidence quality: B
- Benefit: Improve OSAS and accompanying symptoms and sequelae.

- Harm: Expense, time, anxiety; parental sleep disruption; nasal and skin adverse effects; possible midface remodeling; extremely rare serious pressure-related complications, such as pneumothorax; poor adherence.
- Benefits-harms assessment: Preponderance of benefit over harm.
- Value judgments: Panelists believe that CPAP is the most effective treatment of OSAS that persists postoperatively and that the benefits of treatment outweigh the adverse effects. Other treatments (eg, rapid maxillary expansion) may be effective in specially selected patients.
- Role of patient preferences: Other treatments may be effective in specially selected patients.
- Exclusions: Rare patients at increased risk of severe pressure complications.
- Intentional vagueness: None.
- Policy level: Recommendation.

CPAP therapy is delivered by using an electronic device that delivers air at positive pressure via a nasal mask, leading to mechanical stenting of the airway and improved functional residual capacity in the lungs. There is no clear advantage of using bilevel pressure over CPAP.¹⁵ CPAP should be managed by an experienced and skilled clinician with expertise in its use in children. CPAP pressure requirements vary among individuals and change over time; thus, CPAP must be titrated in the sleep laboratory before prescribing the device and periodically readjusted thereafter. Behavioral modification therapy may be required, especially for young children or those with developmental delays. Objective monitoring of adherence, by using the equipment software, is important. If adherence is suboptimal, the clinician should institute measures to improve adherence (such as behavioral modification, or treating side effects of

CPAP) and institute alternative treatments if these measures are ineffective.

Key Action Statement 7: Weight Loss

Clinicians should recommend weight loss in addition to other therapy if a child/adolescent with OSAS is overweight or obese. (Evidence Quality: Grade C, Recommendation Strength: Recommendation.)

Evidence Profile KAS 7

- Aggregate evidence quality: C
- Benefit: Improve OSAS and accompanying symptoms and sequelae; non-OSAS-related benefits of weight loss.
- Harm: Hard to achieve and maintain weight loss.
- Benefits-harms assessment: Preponderance of benefit over harm.
- Value judgments: The panel agreed that weight loss is beneficial for both OSAS and other health issues, but clinical experience suggests that weight loss is difficult to achieve and maintain, and even effective weight loss regimens take time; therefore, additional treatment is required in the interim.
- Role of patient preferences: Strong role for patient and family preference regarding nutrition and exercise.
- Exclusions: None.
- Intentional vagueness: None.
- Strength: Recommendation.

Weight loss has been shown to improve OSAS,^{16,17} although the degree of weight loss required has not been determined. Because weight loss is a slow and unreliable process, other treatment modalities (such as adenotonsillectomy or CPAP therapy) should be instituted until sufficient weight loss has been achieved and maintained.

Key Action Statement 8: Intranasal Corticosteroids

Clinicians may prescribe topical intranasal corticosteroids for children with mild OSAS in whom adenotonsillectomy is contraindicated or for children with mild post-operative OSAS. (Evidence Quality: Grade B, Recommendation Strength: Option.)

Evidence Profile KAS 8

- Aggregate evidence quality: B
- Benefit: Improves mild OSAS and accompanying symptoms and sequelae.
- Harm: Some subjects may not have an adequate response. It is not known whether therapeutic effect persists long-term; therefore, long-term observation is required. Low risk of steroid-related adverse effects.
- Benefits-harms assessment: Preponderance of benefit over harm.
- Value judgments: The panel agreed that intranasal steroids provide a less invasive treatment than surgery or CPAP and, therefore, may be preferred in some cases despite lower efficacy and lack of data on long-term efficacy.
- Role of patient preferences: Moderate role for patient and family preference if OSAS is mild.
- Exclusions: None.
- Intentional vagueness: None.
- Strength: Option.

Mild OSAS is defined, for this indication, as an AHI <5 per hour, on the basis of studies on intranasal corticosteroids described in the accompanying technical report.² Several studies have shown that the use of intranasal steroids decreases the degree of OSAS; however, although

OSAS improves, residual OSAS may remain. Furthermore, there is individual variability in response to treatment, and long-term studies have not been performed to determine the duration of improvement. Therefore, nasal steroids are not recommended as a first-line therapy. The response to treatment should be measured objectively after a course of treatment of approximately 6 weeks. Because the long-term effect of this treatment is unknown, the clinician should continue to observe the patient for symptoms of recurrence and adverse effects of corticosteroids.

AREAS FOR FUTURE RESEARCH

A detailed list of research recommendations is provided in the accompanying technical report.² There is a great need for further research into the prevalence of OSAS, sequelae of OSAS, best treatment methods, and the role of obesity. In particular, well-controlled, blinded studies, including randomized controlled trials of treatment, are needed to determine the best care for children and adolescents with OSAS.

SUBCOMMITTEE ON OBSTRUCTIVE SLEEP APNEA SYNDROME*

Carole L. Marcus, MBBCh, Chairperson (Sleep Medicine, Pediatric Pulmonologist; Liaison, American Academy of Sleep Medicine; Research Support from Philips Respironics; Affiliated with an academic sleep center; Published research related to OSAS)

Lee J. Brooks, MD (Sleep Medicine, Pediatric Pulmonologist; Liaison, American College of Chest Physicians; No financial conflicts; Affiliated with an academic sleep center; Published research related to OSAS)

Sally Davidson Ward, MD (Sleep Medicine, Pediatric Pulmonologist; No financial conflicts; Affiliated with an academic sleep center; Published research related to OSAS)

Kari A. Draper, MD (General Pediatrician; No conflicts)

David Gozal, MD (Sleep Medicine, Pediatric Pulmonologist; Research support from AstraZeneca; Speaker for Merck Company; Affiliated with an academic sleep center; Published research related to OSAS)

Ann C. Halbower, MD (Sleep Medicine, Pediatric Pulmonologist; Liaison, American Thoracic Society; Research Funding from Resmed; Affiliated with an academic sleep center; Published research related to OSAS)

Jacqueline Jones, MD (Pediatric Otolaryngologist; AAP Section on Otolaryngology-Head and Neck Surgery; Liaison, American Academy of Otolaryngology-Head and Neck Surgery; No financial conflicts; Affiliated with an academic otolaryngologic practice)

Christopher Lehman, MD (Neonatologist, Informatician; No conflicts)

Michael S. Schechter, MD, MPH (Pediatric Pulmonologist; AAP Section on Pediatric Pulmonology; Consultant to Genentech, Inc and Gilead, Inc, not related to Obstructive Sleep Apnea; Research Support from Mpx Pharmaceuticals, Inc, Vertex Pharmaceuticals Incorporated, PTC Therapeutics, Bayer Healthcare, not related to Obstructive Sleep Apnea)

Stephen Sheldon, MD (Sleep Medicine, General Pediatrician; Liaison, National Sleep Foundation; No financial conflicts; Affiliated with an academic sleep center; Published research related to OSAS)

Richard N. Shiffman, MD, MCIS (General Pediatrics, Informatician; No conflicts)

Karen Spruyt, PhD (Clinical Psychologist, Child Neuropsychologist, and Biostatistician/Epidemiologist; No financial conflicts; Affiliated with an academic sleep center)

Oversight from the Steering Committee on Quality Improvement and Management, 2009–2012

STAFF

Caryn Davidson, MA

*Areas of expertise are shown in parentheses after each name.

ACKNOWLEDGMENTS

The committee thanks Jason Caboot, June Chan, Mary Currie, Fiona Healy, Maureen Josephson, Sofia Konstantinopoulou, H. Madan Kumar, Roberta Leu, Darius Loghmanee, Rajeev Bhatia, Argyri Petrocheilou, Harsha Vardhan, and Colleen Walsh for assisting with evidence extraction.

REFERENCES

1. Section on Pediatric Pulmonology, Subcommittee on Obstructive Sleep Apnea Syndrome. American Academy of Pediatrics. Clinical

practice guideline: diagnosis and management of childhood obstructive sleep apnea syndrome. *Pediatrics*. 2002;109(4):704–712

2. Marcus CL, Brooks LJ, Davidson C, et al; American Academy of Pediatrics, Subcommittee on Obstructive Sleep Apnea

- Syndrome. Technical report: diagnosis and management of childhood obstructive sleep apnea syndrome. *Pediatrics*. 2012; 130(3):In press
3. American Academy of Pediatrics Steering Committee on Quality Improvement and Management. Classifying recommendations for clinical practice guidelines. *Pediatrics*. 2004;114(3):874–877
 4. American Thoracic Society. Standards and indications for cardiopulmonary sleep studies in children. *Am J Respir Crit Care Med*. 1996;153(2):866–878
 5. Bixler EO, Vgontzas AN, Lin HM, et al. Sleep disordered breathing in children in a general population sample: prevalence and risk factors. *Sleep*. 2009;32(6):731–736
 6. Li AM, So HK, Au CT, et al. Epidemiology of obstructive sleep apnoea syndrome in Chinese children: a two-phase community study. *Thorax*. 2010;65(11):991–997
 7. O'Brien LM, Holbrook CR, Mervis CB, et al. Sleep and neurobehavioral characteristics of 5- to 7-year-old children with parentally reported symptoms of attention-deficit/hyperactivity disorder. *Pediatrics*. 2003; 111(3):554–563
 8. Himes JH, Dietz WH; The Expert Committee on Clinical Guidelines for Overweight in Adolescent Preventive Services. Guidelines for overweight in adolescent preventive services: recommendations from an expert committee. *Am J Clin Nutr*. 1994;59(2):307–316
 9. Mitchell RB. Adenotonsillectomy for obstructive sleep apnea in children: outcome evaluated by pre- and postoperative polysomnography. *Laryngoscope*. 2007;117(10):1844–1854
 10. Suen JS, Arnold JE, Brooks LJ. Adenotonsillectomy for treatment of obstructive sleep apnea in children. *Arch Otolaryngol Head Neck Surg*. 1995;121(5):525–530
 11. Nieminen P, Tolonen U, Löppönen H. Snoring and obstructive sleep apnea in children: a 6-month follow-up study. *Arch Otolaryngol Head Neck Surg*. 2000;126(4):481–486
 12. Blunden S, Lushington K, Lorenzen B, Wong J, Balendran R, Kennedy D. Symptoms of sleep breathing disorders in children are underreported by parents at general practice visits. *Sleep Breath*. 2003;7(4):167–176
 13. Apostolidou MT, Alexopoulos EI, Chaidas K, et al. Obesity and persisting sleep apnea after adenotonsillectomy in Greek children. *Chest*. 2008;134(6):1149–1155
 14. Mitchell RB, Kelly J. Outcome of adenotonsillectomy for severe obstructive sleep apnea in children. *Int J Pediatr Otorhinolaryngol*. 2004;68(11):1375–1379
 15. Marcus CL, Rosen G, Ward SL, et al. Adherence to and effectiveness of positive airway pressure therapy in children with obstructive sleep apnea. *Pediatrics*. 2006; 117(3). Available at: www.pediatrics.org/cgi/content/full/117/3/e442
 16. Verhulst SL, Franckx H, Van Gaal L, De Backer W, Desager K. The effect of weight loss on sleep-disordered breathing in obese teenagers. *Obesity (Silver Spring)*. 2009;17(6):1178–1183
 17. Kalra M, Inge T. Effect of bariatric surgery on obstructive sleep apnoea in adolescents. *Paediatr Respir Rev*. 2006;7(4):260–267

Diagnosis and Management of Childhood Obstructive Sleep Apnea Syndrome

Carole L. Marcus, Lee Jay Brooks, Kari A. Draper, David Gozal, Ann Carol Halbower, Jacqueline Jones, Michael S. Schechter, Stephen Howard Sheldon, Karen Spruyt, Sally Davidson Ward, Christopher Lehmann and Richard N. Shiffman

Pediatrics; originally published online August 27, 2012;

DOI: 10.1542/peds.2012-1671

Updated Information & Services

including high resolution figures, can be found at:
<http://pediatrics.aappublications.org/content/early/2012/08/22/peds.2012-1671>

Citations

This article has been cited by 18 HighWire-hosted articles:
<http://pediatrics.aappublications.org/content/early/2012/08/22/peds.2012-1671#related-urls>

Permissions & Licensing

Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
<http://pediatrics.aappublications.org/site/misc/Permissions.xhtml>

Reprints

Information about ordering reprints can be found online:
<http://pediatrics.aappublications.org/site/misc/reprints.xhtml>

PEDIATRICS is the official journal of the American Academy of Pediatrics. A monthly publication, it has been published continuously since 1948. PEDIATRICS is owned, published, and trademarked by the American Academy of Pediatrics, 141 Northwest Point Boulevard, Elk Grove Village, Illinois, 60007. Copyright © 2012 by the American Academy of Pediatrics. All rights reserved. Print ISSN: 0031-4005. Online ISSN: 1098-4275.

American Academy of Pediatrics

DEDICATED TO THE HEALTH OF ALL CHILDREN™



Retained tympanostomy tubes removal for ICD-10

Question: How should the retained tympanostomy tube guideline be modified for the ICD-10 Prioritized List?

Question source: Holly Jo Hodges, Trillium CCO; HERC Staff

Issue: In August 2014 HERC approved a new guideline note for retained tympanostomy tubes.

Guideline Note XXX Retained tympanostomy tubes

Lines 178, 308, 405, 418, 502

Removal of retained tympanostomy tubes under anesthesia, if indicated (cpt code 69424 *Ventilating tube removal requiring general anesthesia*) or as part of an office visit, are intended to be covered for Line 502 diagnoses with the Line 405 icd-9 code 385.83 *Retained foreign body of middle ear*.

This needs to be applied to the ICD-10 list, however, it does not translate well.

The equivalent Line on the ICD-10 List is 379 CHOLESTEATOMA; INFECTIONS OF THE PINNA but does not have a comparable code to retained foreign body of middle ear. The code it translates to is: *H74.8Xx Other specified disorders of middle ear and mastoid, unspecified ear*

ICD-10 Prioritized List Status

Cpt code 69424 *Ventilating tube removal requiring general anesthesia* is located on the following lines:

Line	Condition	Treatment
174	ACUTE MASTOIDITIS	MASTOIDECTOMY, MEDICAL THERAPY
290	COMPLICATIONS OF A PROCEDURE ALWAYS REQUIRING TREATMENT	MEDICAL AND SURGICAL TREATMENT
317	HEARING LOSS - AGE 5 OR UNDER	MEDICAL THERAPY INCLUDING HEARING AIDS
379	CHOLESTEATOMA; INFECTIONS OF THE PINNA	MEDICAL AND SURGICAL TREATMENT
394	ACUTE OTITIS MEDIA	MEDICAL AND SURGICAL TREATMENT
481	CHRONIC OTITIS MEDIA	PE TUBES/ADENOIDECTOMY/TYMPAN OPLASTY, MEDICAL THERAPY

ICD-10 Code	Code Description	Line
H74.8x1	Other specified disorders of right middle ear and mastoid	481 CHRONIC OTITIS MEDIA
H74.8x2	Other specified disorders of left middle ear and mastoid	481 CHRONIC

Retained tympanostomy tubes removal for ICD-10

ICD-10 Code	Code Description	Line
		OTITIS MEDIA
H74.8x3	Other specified disorders of middle ear and mastoid, bilateral	481 CHRONIC OTITIS MEDIA
H74.8x9	Other specified disorders of middle ear and mastoid, unspecified ear	481 CHRONIC OTITIS MEDIA

GUIDELINE NOTE 29, TYMPANOSTOMY TUBES IN ACUTE OTITIS MEDIA

Line 394

Tympanostomy tubes (69436) are only included on this line as treatment for 1) recurrent acute otitis media (three or more episodes in six months or four or more episodes in one year) that fail appropriate medical management, 2) for patients who fail medical treatment secondary to multiple drug allergies or who fail two or more consecutive courses of antibiotics, or 3) complicating conditions (immunocompromised host, meningitis by lumbar puncture, acute mastoiditis, sigmoid sinus/jugular vein thrombosis by CT/MRI/MRA, cranial nerve paralysis, sudden onset dizziness/vertigo, need for middle ear culture, labyrinthitis, or brain abscess). Patients with craniofacial anomalies, Down's syndrome, cleft palate, and patients with speech and language delay may be considered for tympanostomy if unresponsive to appropriate medical treatment or having recurring infections (without needing to meet the strict "recurrent" definition above).

GUIDELINE NOTE 51, CHRONIC OTITIS MEDIA WITH EFFUSION

Line 481

Antibiotic and other medication therapy (including antihistamines, decongestants and nasal steroids) are not indicated for children with chronic otitis media with effusion (OME) (without another appropriate diagnosis).

There should be a 3 to 6 month watchful waiting period after diagnosis of otitis media with effusion, and if documented hearing loss is greater than or equal to 25dB in the better hearing ear, tympanostomy surgery may be indicated given short but not long-term improvement in hearing. Formal audiometry is indicated for children with chronic OME present for 3 months or longer. Children with language delay, learning problems, or significant hearing loss should have hearing testing upon diagnosis. Children with chronic OME who are not at risk for language or developmental delay should be reexamined at 3- to 6-month intervals until the effusion is no longer present, significant hearing loss is identified, or structural abnormalities of the eardrum or middle ear are suspected.

For the child who has had chronic OME and who has a hearing deficiency in the better-hearing ear of 25 dB or greater, myringotomy with tube insertion is recommended after a total of 4 to 6 months of effusion with a documented hearing deficit.

Adenoidectomy is not indicated at the time of first pressure equalization tube insertion. It may be indicated in children over 3 years who are having their second set of tubes.

Tube insertion should be covered for patients with craniofacial anomalies, Down's syndrome, cleft palate and patients with speech and language delay along with co-morbid hearing loss.

HERC Staff Assessment

H74.8xX could be added to Line 379 to be consistent. However, this would be adding a new diagnosis to a currently unrelated line. More simply, retained tympanostomy tubes can just be viewed as a complication, since these usually fall out on their own. Removal of tubes is currently on the Complications ALWAYS requiring treatment line (290) which generally has much higher acuity issues on it. So, placing retained tympanostomy tubes

Retained tympanostomy tubes removal for ICD-10

and removal of these tubes to the Complications USUALLY Requiring Treatment line is logical.

HERC Staff Recommendations:

- 1) Remove 69424 *Ventilating tube removal requiring general anesthesia* from all lines. Place only on Line 427 COMPLICATIONS OF A PROCEDURE USUALLY REQUIRING TREATMENT
- 2) For ICD-9 code 385.83 RETAINED FOREIGN BODY OF MIDDLE EAR
 - a. Remove from line 379 CHOLESTEATOMA; INFECTIONS OF THE PINNA
 - b. Add to Line 427
- 3) For ICD-10 code H74.8xX *Other specified disorders of middle ear and mastoid* Add to Line 427
- 4) Delete the guideline note on RETAINED TYMPANOSTOMY TUBES approved at the August meeting and instead modify guideline notes 29 and 51 as follows:

GUIDELINE NOTE 29, TYMPANOSTOMY TUBES IN ACUTE OTITIS MEDIA

Line 394

Tympanostomy tubes (69436) are only included on this line as treatment for 1) recurrent acute otitis media (three or more episodes in six months or four or more episodes in one year) that fail appropriate medical management, 2) for patients who fail medical treatment secondary to multiple drug allergies or who fail two or more consecutive courses of antibiotics, or 3) complicating conditions (immunocompromised host, meningitis by lumbar puncture, acute mastoiditis, sigmoid sinus/jugular vein thrombosis by CT/MRI/MRA, cranial nerve paralysis, sudden onset dizziness/vertigo, need for middle ear culture, labyrinthitis, or brain abscess). Patients with craniofacial anomalies, Down's syndrome, cleft palate, and patients with speech and language delay may be considered for tympanostomy if unresponsive to appropriate medical treatment or having recurring infections (without needing to meet the strict "recurrent" definition above).

[Removal of retained tympanostomy tubes requiring anesthesia \(CPT code 69424\) or as an office visit, is included on line 427 as a complication, pairing with 385.83/ H74.8xX.](#)

GUIDELINE NOTE 51, CHRONIC OTITIS MEDIA WITH EFFUSION

Line 481

Antibiotic and other medication therapy (including antihistamines, decongestants and nasal steroids) are not indicated for children with chronic otitis media with effusion (OME) (without another appropriate diagnosis).

Retained tympanostomy tubes removal for ICD-10

There should be a 3 to 6 month watchful waiting period after diagnosis of otitis media with effusion, and if documented hearing loss is greater than or equal to 25dB in the better hearing ear, tympanostomy surgery may be indicated given short but not long-term improvement in hearing. Formal audiometry is indicated for children with chronic OME present for 3 months or longer. Children with language delay, learning problems, or significant hearing loss should have hearing testing upon diagnosis. Children with chronic OME who are not at risk for language or developmental delay should be reexamined at 3- to 6-month intervals until the effusion is no longer present, significant hearing loss is identified, or structural abnormalities of the eardrum or middle ear are suspected.

For the child who has had chronic OME and who has a hearing deficiency in the better-hearing ear of 25 dB or greater, myringotomy with tube insertion is recommended after a total of 4 to 6 months of effusion with a documented hearing deficit.

Adenoidectomy is not indicated at the time of first pressure equalization tube insertion. It may be indicated in children over 3 years who are having their second set of tubes.

Tube insertion should be covered for patients with craniofacial anomalies, Down's syndrome, cleft palate and patients with speech and language delay along with co-morbid hearing loss.

[Removal of retained tympanostomy tubes requiring anesthesia \(CPT code 69424\) or as an office visit, is included on line 427 as a complication, pairing with 385.83/ H74.8xX.](#)

Breastfeeding supplies and support

Question: Should the guideline on coverage of breastfeeding support and supplies be clarified?

Question Source: Breastfeeding Coalition of Oregon
Christine Seals, CCO Medical Director
Helen Bellanca, Maternal Child Family Program Director, Health Share

Issue summary: There are concerns about inadequate or inconsistent coverage of both breastpumps and lactation support among OHP recipients. The January 1, 2015 Prioritized List refers to the HRSA guidelines to be covered on Line 3. HRSA specifically requires costs for renting breastfeeding equipment and comprehensive lactation support to be covered. The federal ACA requires coverage of breast pumps for lactating mothers and lactation counseling support for commercial plans.

Both WIC and OHP currently provide breastpumps. Sometimes OHP CCOs are providing pumps as first line, and sometimes mothers are being referred to WIC if they need pumps. There are also concerns that inadequate WIC penetration may mean that

There is currently no OHP coverage of lactation visits outside of the immediate postpartum hospital stay where it is bundled within the delivery hospital stay.

There are implementation concerns about paying for lactation consultants separately based on the fact that these are often nonlicensed providers.

From WIC, Bonnie Ranno

Staff in WIC are excited about the inclusion of lactation services on HERC's prioritized list. There is such good evidence that the provision of lactation support, counseling, and equipment can increase breastfeeding exclusivity and duration, which then translates to significant health care savings and improved health for mothers and children.

Our WIC breastfeeding team has gathered some information to help guide your writing of the HERC Guideline on Lactation Services. Please see the following attached documents:

- Recommendations and evidence base for the provision of lactation services in Oregon, with references
- Examples of what other states are providing in regards to Medicaid coverage of lactation services

You had also asked about whether WIC would be able to change their policies regarding the provision of breast pumps for women returning to work or school. While WIC's funding for breast pumps is uncertain from year to year, we have been able to provide pumps for a number of WIC participants, and expect that we will be able to continue to do so, especially with some shifting from hospital pumps to personal pumps. It appears unlikely, however, that WIC's guidelines for electric pump distribution will broaden to include women who are not exclusively breastfeeding or those who are only away from their babies for a few hours. Those mothers would still qualify for a manual pump, though. I will also attach a FAQs document re Oregon WIC's breast pump program, which provides additional information.

Breastfeeding supplies and support

In closing, we just want to reiterate that the HERC guideline will help to correct a real problem. CCOs have experienced a lot of confusion about lactation coverage, and clear guidance is needed. Putting lactation services on the prioritized list will help Oregon to

- Correct the equity issues that exist across the state in regard to lactation services
- Save money by preventing current and future health problems of mothers and babies
- Implement evidence-based and best practices pertaining to breastfeeding
- Fulfill Oregon Medicaid's requirement of complying with the Affordable Care Act

Prioritized List status (as of January 1, 2015)

Line 3

Included on this line are the following preventive services:

1. US Preventive Services Task Force (USPSTF) "A" and "B" Recommendations (as of May 2012):
<http://www.uspreventiveservicestaskforce.org/uspstf/uspsabrecs.htm>
2. American Academy of Pediatrics (AAP) Bright Futures Guidelines (published 2008):
<http://brightfutures.aap.org/pdfs/aap%20bright%20futures%20periodicity%20sched%20101107.pdf>
3. Health Resources and Services Administration (HRSA) Women's Preventive Services - Required Health Plan Coverage Guidelines: (approved with Affordable Care Act on March 23, 2010)
<http://www.hrsa.gov/womensguidelines/>
4. Immunizations as recommended by the Advisory Committee on Immunization Practices (ACIP) and approved for the Oregon Immunization Program:
<http://public.health.oregon.gov/PreventionWellness/VaccinesImmunization/ImmunizationProviderResources/Documents/DMApvactable.pdf>

HRSA Guideline (referred to in #3 above)

Type of Preventive Service	HHS Guideline for Health Insurance Coverage	Frequency
Breastfeeding support, supplies, and counseling.	Comprehensive lactation support and counseling, by a trained provider during pregnancy and/or in the postpartum period, and costs for renting breastfeeding equipment.	In conjunction with each birth.

USPSTF, 2008

Recommendation: The USPSTF recommends interventions during pregnancy and after birth to promote and support breastfeeding (Grade B recommendation).

Breastfeeding supplies and support

From the Healthcare.gov website describing benefits:

<https://www.healthcare.gov/what-are-my-breastfeeding-benefits/>

The health care law requires most health insurance plans to provide breastfeeding equipment and counseling for pregnant and nursing women.

Health insurance plans must provide breastfeeding support, counseling, and equipment for the duration of breastfeeding. These services may be provided before and after you have your baby.

These rules apply to Health Insurance Marketplace plans and all other health insurance plans, except for grandfathered plans.

Coverage of breast pumps

Your health insurance plan must cover the cost of a breast pump – and may offer to cover either a rental or a new one for you to keep.

Your plan may provide guidance on whether the covered pump is manual or electric, how long the coverage of a rented pump lasts, and when they'll provide the pump (before or after you have the baby).

Code	Code Description	Line placement
99401	Preventive medicine counseling and/or risk factor reduction intervention(s) provided to an individual (separate procedure); approximately 15 minutes	650 lines.
99402	Preventive medicine counseling and/or risk factor reduction intervention(s) provided to an individual (separate procedure); approximately 15 minutes	650 lines.
99403	Preventive medicine counseling and/or risk factor reduction intervention(s) provided to an individual (separate procedure); approximately 45 minutes	650 lines
99404	Preventive medicine counseling and/or risk factor reduction intervention(s) provided to an individual (separate procedure); approximately 60 minutes	650 lines
99411	Preventive medicine counseling and/or risk factor reduction intervention(s) provided to individuals in a group setting (separate procedure); approximately 30 minutes	650 lines
99412	Preventive medicine counseling and/or risk factor reduction intervention(s) provided to individuals in a group setting (separate procedure); approximately 60 minutes	650 lines
99078	Physician or other qualified health care professional qualified by education, training, licensure/regulation (when applicable) educational services rendered to patients in a group setting (eg, prenatal, obesity, or diabetic instructions)	650 lines
S9443	Lactation classes, non-physician provider, per session	DMAP Ancillary Codes File
V24.1	Postpartum care and examination of lactating mother	1 PREGNANCY

Breastfeeding supplies and support

Code	Code Description	Line placement
Z39.1	Encounter for care and examination of lactating mother	1 PREGNANCY
Z39.2	Encounter for routine postpartum follow-up	1,6

WIC criteria

Does the mother or the baby have a **medical need** for a pump?

Possible examples include:

- ☐ Infant is unable to nurse due to prematurity or another health problem, or
- ☐ Mother needs to increase milk production (multiple birth, previous breast surgery)

If yes, Loan Multi-User Double Electric Breast Pump

Is the mother going to work or school more than 20 hrs/wk **and**,

- ☐ Is the infant at least 4 weeks old, **and**
- ☐ Is the infant exclusively breastfeeding, **and**
- ☐ Are mother and baby not having any problems with breastfeeding, **and**
- ☐ Is mother planning to breastfeed exclusively for at least four to six months and not expecting any formula from WIC?

If yes, personal double electric pump offered. If no, hand pump recommended.

DMAP rules

410-122-0250 Breast Pumps

(1) Indications and limitations of coverage and medical appropriateness:

(a) The Division of Medical Assistance Programs (Division) may cover electric breast pumps for any of the following conditions:

(A) Medical appropriateness for infant:

- (i) Pre-term;
- (ii) Term and hospitalized beyond five days;
- (iii) Separated from mother for an undetermined length of time;
- (iv) Cleft palate or cleft lip;
- (v) Cranial-facial abnormalities;
- (vi) Inability to suck adequately;
- (vii) Re-hospitalized for longer than two days;
- (viii) Failure to thrive;

(B) Medical appropriateness for mother:

- (i) Breast abscess;
- (ii) Mastitis;
- (iii) Hospitalized due to illness or surgery (for short-term use to maintain lactation);
- (iv) Short-term treatment with medications that may be transmitted to the infant;
- (v) A hand pump or manual expression has been tried for one week without success in mothers with established milk supply;

(b) Documentation that transition to breast feeding started as soon as the infant was stable enough to begin breast feeding;

(c) Use E1399 for an electric breast pump starter kit for single or double pumping;

Breastfeeding supplies and support

(d) An electric breast pump starter kit will be reimbursed separately from the breast pump rental; Durable Medical Equipment, Prosthetics, Orthotics and Supplies

Breastfeeding supplies and support

- (e) Electric breast pump rental cannot exceed 60 days,
 - (f) An electric breast pump may only be purchased when cost effective for one of the following conditions:
 - (i) Cleft palate or cleft lip;
 - (ii) Cranial-facial abnormalities;
 - (iii) Inability to suck adequately;
 - (iv) Infant is separated from mother for an undetermined length of time;
 - (g) Electric breast pump rental charges apply to the purchase price;
 - (h) The following services are not covered:
 - (i) Accessories;
 - (ii) An electric breast pump for the comfort and convenience of the mother;
 - (iii) Supplemental Nutrition System (SNS);
 - (iv) Heavy duty, hospital grade breast pumps;
 - (v) Replacement parts.
- (2) Documentation requirements:
 - (a) For services that require prior authorization (PA): Submit documentation for review which supports conditions of coverage as specified in this rule are met;
 - (b) For services that do not require PA: Medical records which support conditions of coverage as specified in this rule are met must be on file with the durable medical equipment, prosthetics, orthotics and supplies (DMEPOS) provider and made available to the Division on request.
- (3) Procedure Codes:
 - (a) E0602 – Breast pump, manual, any type – the Division will purchase;
 - (b) E0603 – Breast pump, electric (AC and/or DC), any type:
- (A) The Division will purchase or rent on a monthly basis;
- (B) PA required

Breastfeeding supplies and support

- (c) E1399 – Electric breast pump starter kit;
- (A) The Division will purchase;
- (B) PA required.

Evidence

USPSTF, 2008

- 1) **The USPSTF recommends interventions during pregnancy and after birth to promote and support breastfeeding.**
Grade: [B Recommendation](#).
- 2) **Strategies that include prenatal and postnatal components are most effective and strategies should include:**
 - a. **formal breastfeeding education for mothers**
 - b. **direct support of mothers during breastfeeding**
 - c. **training of primary care staff about breastfeeding and techniques for breastfeeding support**
 - d. **peer support**
- 3) **system level interventions encouraged**

AHRQ, 2007

1. Health Technology Assessment
2. Assessment of impacts of breastfeeding on maternal and infant health in developed countries
3. 29 systematic reviews or metaanalysis included, 43 unique primary studies as well
4. In infants:
 - a. Found that breastfeeding significantly reduces the risk of:
 - i. Acute otitis media
 - ii. Severe (i.e. requiring hospitalization) lower respiratory tract diseases
 - iii. Non-specific gastroenteritis
 - iv. Necrotizing enterocolitis
 - v. Atopic dermatitis,
 - vi. Asthma (young children)
 - vii. Obesity
 - viii. Type 1 and 2 diabetes
 - ix. Childhood leukemia
 - x. Sudden infant death syndrome (SIDS)
 - b. Found no clear association with
 - i. Cognitive performance
 - ii. Mortality (in developed countries)
 - iii. Cardiovascular outcomes
5. In women,
 - a. A history of lactation is associated with:
 - b. Reduced risk of type 2 diabetes
 - c. Reduced risk of breast cancer
 - d. Reduced risk of ovarian cancer.
 - e. Protective effect against postpartum depression
 - f. No clear relationship was found between a history of lactation and
 - i. Osteoporosis
 - ii. Return to pre-pregnancy weight (negligible) and postpartum weight loss

Breastfeeding supplies and support

WHO 2013

- 1) Systematic review, update of 2007 systematic review
- 2) Most studies are observational, and potentially subject to selection bias and confounding. Most studies were conducted in high income countries.
- 3) 2 RCTs exist, one to formula versus breastfeeding and the other to breastfeeding counseling. Both demonstrated a higher IQ in breastfed infants but no significant impacts on other health outcomes. Maternal IQ was accounted for.
- 4) Concluding differences (pp 67-68):

TABLE 10.1

Pooled effects for each outcome, from all studies and from those deemed to be of higher scientific quality

Outcome	Pooled effect (95% confidence interval)	
	All studies	High-quality studies ^a
Mean total blood cholesterol (mmol/L)	-0.01 (-0.05; 0.02)	0.00 (-0.02; 0.02)
Mean systolic blood pressure (mmHg)	-1.02 (-1.45; -0.59)	-0.71 (-1.24; -0.19)
Mean diastolic blood pressure (mmHg)	-0.37 (-0.71; -0.04)	-0.27 (-0.64; 0.09)
Odds ratio of type-2 diabetes	0.66 (0.49-0.89)	Not estimated
Odds ratio of overweight/obesity	0.76 (0.71; 0.81)	0.88 (0.83; 0.93)
Mean performance in intelligence test (points)	3.45 (1.92-4.98)	2.19 (0.89-3.50)

^a High-quality studies include those with larger sample sizes and adjustment for confounding variables relevant to each outcome (see individual chapters for further details).

Johns, 2013

1. Systematic review of breast milk expression, prevalence and outcomes
2. 7 papers between 1999 and 2012 identified
3. Breastmilk feeding only at the breast for the duration of breastfeeding is rare in developed countries
4. Many reasons for breastpumping:
 - a. Difficulty establishing breastfeeding
 - b. Mothers with premature or low birth weight infants
 - c. Mothers who are unwell
 - a. Elevated BMI – expression may increase duration of breastfeeding among obese mothers
 - b. Mastitis, breast engorgement, nipple pain
 - c. Concerns about over or under supply
 - d. Return to paid work
 - e. Cultural factors and also embarrassment about feeding in public
 - f. Ability to have others perform the “task” of feeding the infant
5. Impact on breastfeeding outcomes
 - a. Contradictory results
 - b. Impact on maternal mental health contested

Cochrane 2012

1. Systematic review of duration of breastfeeding, comparing at least 6 months exclusive to 3-4 months mixed breastfeeding

Breastfeeding supplies and support

2. Found less gastrointestinal infection, and possibly respiratory infection among those exclusively versus partially breastfed. Lactational amenorrhea and weight loss may be maternal benefits
3. Supported recommendation of WHO for exclusive breastfeeding x 6 months

Guidance from CMS

http://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Quality-of-Care/Downloads/Lactation_Services_IssueBrief_01102012.pdf

State Medicaid survey in 2008 found:

- 25 of the responding States covered breastfeeding education services
- 15 of the responding States covered individual lactation consultations
- 31 of the responding States covered equipment rentals.

“Due to the multiple health benefits associated with breastfeeding, CMS encourages States to go beyond the requirement of solely coordinating and referring enrollees to the Special Supplemental Food Program for Women, Infants, and Children (WIC) (established in 42 C.F.R. § 431.635) and include lactation services as separately reimbursed pregnancy-related services.

States may use the following Medicaid coverage categories to reimburse lactation services:

- Inpatient hospital services (other than services in an institution for mental disease), per Social Security Act (SSA) § 1905(a)(1);
- Outpatient hospital services, per SSA § 1905(a)(2)(A) and 42 C.F.R. § 440.10;
- Early and periodic screening, diagnostic, and treatment services for individuals who are eligible under the plan and are under the age of 21, per SSA § 1905(a)(4)(B);
- Physicians' services furnished by a physician under the physician's supervision, whether furnished in the office, the patient's home, a hospital, or a nursing facility, or elsewhere, per SSA § 1905(a)(5)(A);
- Services furnished by a nurse-midwife, which the nurse-midwife is legally authorized to perform under State law, per SSA § 1905(a)(17);
- Freestanding birth center services, per SSA § 1905(a)(28); and
- Services furnished by nurse practitioner “

They recommend to:

- Eliminate variation of coverage among Medicaid managed care plans and encourage managed care contractors to provide breastfeeding education, either by means of referrals to Special Supplemental Food Program for Women, Infants, and Children (WIC) or direct provision of the services through its provider network. WIC provides breastfeeding women with lactation support by means of counseling, the provision of breastfeeding educational materials, the option to receive an enhanced food package, follow-up support through peer counselors, and the provision of breast pumps or other equipment to help reinforce a healthy and successful breastfeeding routine.

WIC

1. Policy paper (see WIC - HERC Guideline Information, with References)
2. Recommend strongly covering lactation support services
3. Draw attention to the IBCLC certification as well as doulas and traditional health workers as potential sources of lactation assessment and support and peer supports
4. Provide examples of other states successfully implementing payment (see WIC - State Examples of Lactation Support payments)

Breastfeeding supplies and support

Cost savings analysis, North Carolina Medicaid

North Carolina Medicaid performed a cost analysis

1. If North Carolina increased breastfeeding rates up to Healthy People 2020 goals then:
 - a. 14-18 infant deaths would be prevented
 - b. With 75% getting some lactation support, estimate \$4.77 million in costs, \$7 million saved
 - c. Cost savings

Condition	NC Medicaid Cases Averted	Estimated Cost Savings
Lower Respiratory Tract Infections	500	\$2.5 Million
Gastroenteritis	6,000	\$2.5 Million
Necrotizing Enterocolitis	10	\$2.1 Million
Total Cost		\$7.1 Million

Professional associations

United states breastfeeding committee makes the following “model payer recommendations”

Breastfeeding Support & Counseling Services Benefit Guidelines

	Coverage	Permitted Providers	Coverage Criteria	Member Costs, Exclusions, Maximums
Education & Support				
Prenatal/postpartum breastfeeding classes	Covered benefit for one class series in the prenatal period and one class series up to 12 months postpartum	Approved lactation care providers, ⁸ licensed providers (such as MD, APRN, DO, PA), contracted network providers, RN certified childbirth/breastfeeding educators, and Registered Dietitians (RD, RDN, LDN) ⁹	No referral or prior authorization required	No cost to member Prenatal and postpartum series up to a total of 18 class weeks
Breastfeeding support groups	Covered benefit for one monthly meeting up to a maximum of 12 meetings	Approved lactation care providers, licensed providers (such as MD, APRN, DO, PA), contracted network providers, RN certified childbirth/breastfeeding educators, and Registered Dietitians (RD, RDN, LDN)	No referral or prior authorization required	No cost to member One monthly meeting up to a total of 12 meetings
Lactation Counseling				
In-patient hospital	Covered benefit included as part of inpatient stay for facility employed providers; separately payable service (with place of service code) for non-employee in-network providers	Network hospitals, employed approved lactation care and/or licensed providers; in-network non-employee providers	No referral or prior authorization required	No cost to member
Outpatient hospital, home, provider office/clinic	Covered benefit for 6 outpatient lactation consultation visits per birth up to 2 hours per visit	Approved lactation care providers, licensed providers (such as MD, APRN, DO, PA); non-licensed providers, RN, and RD, RDN, LDN may bill as ‘incident-to’ physician services ¹⁰	No referral or prior authorization required; prior authorization required beyond 6 visits	No cost to member Up to 6 visits per birth, up to 2 hours per visit

HERC Staff Assessment

- 1) Breastfeeding is associated with a large number of improved maternal and infant health outcomes. The HRSA, ACA, USPSTF, and many others recommend comprehensive support to encourage and sustain breastfeeding.
- 2) Pumping - There are situations in which pumping would allow for short term uptake (e.g. medical necessity) and long-term provision of exclusive breastfeeding (e.g. inability of mother to provide breastfeeding directly because of work or school). The Prioritized List indicates breast pumps should be covered by its reference to HRSA Guidelines. Major federal and national organizations universally recommend coverage.

Breastfeeding supplies and support

- a. There is inconsistency across CCOs and DMAP about expectations for coverage of breast pumps by OHP versus WIC. Currently, WIC is interested in continuing to provide breastpumps for women returning to work or to school.
- 3) Lactation support – there is not coverage for lactation support aside from what is bundled into the hospital delivery stay
- 4) DMAP is actively revising their rules for breast pump coverage

HERC Staff Recommendation:

- 1. Add codes S9443, V24.1, and Z39.1 to Line 3

S9443	Lactation classes, non-physician provider, per session
V24.1	Postpartum care and examination of lactating mother
Z39.1	Encounter for care and examination of lactating mother

- 2. Add a new guideline

Guideline Note XXX Breastfeeding support and supplies

Line 3

Breastpumps and supplies are covered for postpartum women who are exclusively breastfeeding infants or intending to, and there is maternal or infant medical necessity.

For cases in which there is a medical indication for breast pumps, the pumps should be supplied within 24-48 hours to allow for continued milk production

Lactation support services (including education and counseling by trained providers) are covered for pregnant and postpartum women (*for at least up to 3-6 months*).

Breastfeeding and Maternal and Infant Health Outcomes in Developed Countries

Prepared for:

Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
www.ahrq.gov

Contract No. 290-02-0022

Prepared by:

Tufts-New England Medical Center Evidence-Based Practice Center
Boston, Massachusetts

Investigators

Stanley Ip, M.D., Project Leader
Mei Chung, M.P.H.
Gowri Raman, M.D.
Priscilla Chew, M.P.H.
Nombulelo Magula, M.D.
Deirdre DeVine, M.Litt., Project Manager
Thomas Trikalinos, M.D., Ph.D.
Joseph Lau, M.D., Principal Investigator

AHRQ Publication No. 07-E007
April 2007

This report is based on research conducted by the Tufts-New England Medical Center Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-02-0022). The findings and conclusions in this document are those of the author(s), who are responsible for its content, and do not necessarily represent the views of AHRQ. No statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help clinicians, employers, policymakers, and others make informed decisions about the provision of health care services. This report is intended as a reference and not as a substitute for clinical judgment.

This report may be used, in whole or in part, as the basis for the development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials noted for which further reproduction is prohibited without the specific permission of copyright holders.

Suggested Citation:

Ip S, Chung M, Raman G, Chew P, Magula N, DeVine D, Trikalinos T, Lau J. Breastfeeding and Maternal and Infant Health Outcomes in Developed Countries. Evidence Report/Technology Assessment No. 153 (Prepared by Tufts-New England Medical Center Evidence-based Practice Center, under Contract No. 290-02-0022). AHRQ Publication No. 07-E007. Rockville, MD: Agency for Healthcare Research and Quality. April 2007.

No investigators have any affiliations or financial involvement (e.g., employment, consultancies, honoraria, or stock options, expert testimony, grants, or patents received or pending, or royalties) that conflict with material presented in this report.

Preface

The Agency for Healthcare Research and Quality (AHRQ), through its Evidence-Based Practice Centers (EPCs), sponsors the development of evidence reports and technology assessments to assist public- and private-sector organizations in their efforts to improve the quality of health care in the United States. The Office on Women's Health, Department of Health and Human Services, requested and provided funding for this report. The reports and assessments provide organizations with comprehensive, science-based information on common, costly medical conditions and new health care technologies. The EPCs systematically review the relevant scientific literature on topics assigned to them by AHRQ and conduct additional analyses when appropriate prior to developing their reports and assessments.

To bring the broadest range of experts into the development of evidence reports and health technology assessments, AHRQ encourages the EPCs to form partnerships and enter into collaborations with other medical and research organizations. The EPCs work with these partner organizations to ensure that the evidence reports and technology assessments they produce will become building blocks for health care quality improvement projects throughout the Nation. The reports undergo peer review prior to their release.

AHRQ expects that the EPC evidence reports and technology assessments will inform individual health plans, providers, and purchasers as well as the health care system as a whole by providing important information to help improve health care quality.

We welcome comments on this evidence report. They may be sent by mail to the Task Order Officer named below at: Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850, or by e-mail to epc@ahrq.hhs.gov.

Carolyn M. Clancy, M.D.
Director
Agency for Healthcare Research and Quality

Jean Slutsky, P.A., M.S.P.H.
Director, Center for Outcomes and Evidence
Agency for Healthcare Research and Quality

Suzanne Haynes, Ph.D.
Office on Women's Health
Department of Health and Human Services

Beth Collins Sharp, Ph.D., R.N.
Director, EPC Program
Agency for Healthcare Research and Quality

Ernestine Murray, M.A.S., R.N.
EPC Program Task Order Officer
Agency for Healthcare Research and Quality

Acknowledgments

We acknowledge with appreciation the members of the Technical Expert Panel for their advice and consultation to the Evidence-based Practice Center (EPC) during preparation of this report. In designing the study questions and methodology at the outset of this report, the EPC consulted these technical and content experts. Broad expertise and perspectives are sought. Divergent and conflicted opinions are common and perceived as healthy scientific discourse that results in a thoughtful, relevant systematic review. Therefore, in the end, study questions, design and/or methodologic approaches do not necessarily represent the views of individual technical and content experts.

Technical Expert Panel

David Chelmow, M.D.
Obstetrician-Gynecologist
Tufts-New England Medical Center
Tufts University School of Medicine

Kathryn G. Dewey, Ph.D.
Department of Nutrition
University of California, Davis

Gillman Grave, M.D.
Endocrinology, Nutrition & Growth
National Institute of Child Health and Human
Development

Larry Grummer-Strawn, Ph.D.
National Center for Chronic Disease Prevention
and Health Promotion
Centers for Disease Control and Prevention

Katherine E. Hartmann, M.D., Ph.D.
Obstetrician-Gynecologist
Vanderbilt University
Representative, American College of
Obstetricians and Gynecologists

Suzanne Haynes, Ph.D.
Senior Science Advisor
Office on Women's Health
U.S. Department of Health and Human Services

Jill Janke, RNC, D.NSc.
University of Alaska
Representative, Association of Women's Health,
Obstetrical and Neonatal Nurses

Ronald E. Kleinman, M.D.
Pediatric Gastroenterology and Nutrition
Department of Pediatrics
MassGeneral Hospital for Children

Michael Kramer, M.D.
McGill University Faculty of Medicine
Institute of Human Development, Child and Youth Health
The Montreal Children's Hospital

Katy Lebbing, B.S., IBCLC, RLC
Center for Breastfeeding Information
La Leche League International

Ardythe L. Morrow, Ph.D., M.Sc.
Cincinnati Children's Hospital Medical Center

David Meyers, M.D.
Center for Primary Care, Prevention, and Clinical
Partnerships, (CPPO)
Agency for Healthcare Research and Quality

Barbara L. Philipp, M.D.
The Breastfeeding Center
Boston Medical Center

Lori Feldman-Winter, M.D., M.P.H., IBCLC, FAAP,
FABM
Robert Wood Johnson Medical School
University of Medicine and Dentistry of New Jersey
Representative, AAP and the Section on Breastfeeding

Structured Abstract

Objectives: We reviewed the evidence on the effects of breastfeeding on short- and long-term infant and maternal health outcomes in developed countries.

Data Sources: We searched MEDLINE®, CINAHL, and the Cochrane Library in November of 2005. Supplemental searches on selected outcomes were conducted through May of 2006. We also identified additional studies in bibliographies of selected reviews and by suggestions from technical experts.

Review Methods: We included systematic reviews/meta-analyses, randomized and non-randomized comparative trials, prospective cohort, and case-control studies on the effects of breastfeeding and relevant outcomes published in the English language. Included studies must have a comparative arm of formula feeding or different durations of breastfeeding. Only studies conducted in developed countries were included in the updates of previous systematic reviews. The studies were graded for methodological quality.

Results: We screened over 9,000 abstracts. Forty-three primary studies on infant health outcomes, 43 primary studies on maternal health outcomes, and 29 systematic reviews or meta-analyses that covered approximately 400 individual studies were included in this review. We found that a history of breastfeeding was associated with a reduction in the risk of acute otitis media, non-specific gastroenteritis, severe lower respiratory tract infections, atopic dermatitis, asthma (young children), obesity, type 1 and 2 diabetes, childhood leukemia, sudden infant death syndrome (SIDS), and necrotizing enterocolitis. There was no relationship between breastfeeding in term infants and cognitive performance. The relationship between breastfeeding and cardiovascular diseases was unclear. Similarly, it was also unclear concerning the relationship between breastfeeding and infant mortality in developed countries. For maternal outcomes, a history of lactation was associated with a reduced risk of type 2 diabetes, breast, and ovarian cancer. Early cessation of breastfeeding or not breastfeeding was associated with an increased risk of maternal postpartum depression. There was no relationship between a history of lactation and the risk of osteoporosis. The effect of breastfeeding in mothers on return-to-pre-pregnancy weight was negligible, and the effect of breastfeeding on postpartum weight loss was unclear.

Conclusions: A history of breastfeeding is associated with a reduced risk of many diseases in infants and mothers from developed countries. Because almost all the data in this review were gathered from observational studies, one should not infer causality based on these findings. Also, there is a wide range of quality of the body of evidence across different health outcomes. For future studies, clear subject selection criteria and definition of “exclusive breastfeeding”, reliable collection of feeding data, controlling for important confounders including child-specific factors, and blinded assessment of the outcome measures will help. Sibling analysis provides a method to control for hereditary and household factors that are important in certain outcomes. In addition, cluster randomized controlled studies on the effectiveness of various breastfeeding promotion interventions will provide further opportunity to investigate any disparity in health outcomes as a result of the intervention.

Contents

Executive Summary	1
Evidence Report	11
Chapter 1. Introduction	11
Chapter 2. Methods	13
Overview	13
Key Questions Addressed in This Report	13
Definitions of Breastfeeding in This Report	14
Literature Search Strategy	14
Study Selection	15
Meta-Analysis	16
Grading of Studies Analyzed in This Report	17
Grading of Systematic Reviews/Meta-Analyses.	17
Grading of Individual Primary Studies in Updates and New Reviews	18
Reporting of the Evidence	20
Chapter 3. Results	21
Overview	21
Literature Search Results	23
Organization of Results	24
Part I. Term Infant Outcomes – Relationship of Breastfeeding to Outcomes	27
Acute Otitis Media	27
Atopic Dermatitis	35
Gastrointestinal Infections	37
Lower Respiratory Tract Diseases	40
Asthma	42
Cognitive Development in Terms Infants	50
Obesity	62
Cardiovascular Diseases	68
Type 1 Diabetes	75
Type 2 Diabetes	81
Childhood Leukemia	84
Infant Mortality	90
Sudden Infant Death Syndrome (SIDS)	93
Part II. Preterm Infant Outcomes – Relationship with Human milk feeding	98
Necrotizing Enterocolitis	98
Cognitive Development	103
Part III. Maternal Outcomes – Relationship of Breastfeeding to Outcomes	111
Return to Pre-pregnancy Weight or Postpartum Weight Change	111
Type 2 Diabetes	119

Osteoporosis.....	124
Postpartum Depression	130
Breast Cancer	136
Ovarian Cancer	141
Other Research.....	155
Chapter 4. Discussion	157
Future Research	162
References and Included Studies	165
Acronyms and Abbreviations	185

Figures

Figure 1. The relationship between breastfeeding and health outcomes in term infants – meta-analysis results	21
Figure 2. The relationship between <i>exclusive</i> breastfeeding and health outcomes in term infants - meta-analysis results	22
Figure 3. The relationship between breastfeeding and maternal outcomes - meta-analysis results	23
Figure 4. Primary studies available to assess the relationship between breastfeeding and infant health outcomes	25
Figure 5. Primary studies available to assess the relationship between breastfeeding and maternal health outcomes	26
Figure 6. Meta-analysis of the association between breastfeeding and the risk of AOM compared to exclusive bottle-feeding in cohort studies.....	30
Figure 7. Meta-Analysis of prospective cohort studies of the association between asthma risk and breastfeeding ≥ 3 months for children with positive family history of asthma or atopy (excluding Wright 2001).....	44
Figure 8. Meta-Analysis of prospective cohort studies of the association between asthma risk and breastfeeding ≥ 3 months for children with positive family history of asthma or atopy (including Wright 2001)	45
Figure 9. Meta-Analysis of prospective cohort studies of the association between asthma risk and breastfeeding ≥ 3 months for children without family history of asthma or atopy.....	45
Figure 10. Random effects model of summary estimate evaluating the association of breastfeeding and SIDS.....	95
Figure 11. Meta-Analysis of four RCTs on the effects of breast milk feeding and NEC in preterm infants	100
Figure 12. Meta-Analysis of case-control studies on the relationship between breastfeeding and maternal ovarian cancer risk: ever breastfed versus never breastfed.....	143
Figure 13. Meta-Analysis of case-control studies on the relationship between breastfeeding and maternal ovarian cancer risk: breastfed less than 12 months (cumulative duration) versus never breastfed.....	144

Figure 14. Meta-Analysis of case-control studies on the relationship between breastfeeding and maternal ovarian cancer risk: breastfed at least 12 months (cumulative duration) versus never breastfed.....	145

Tables

Table 1. Summary of systematic review/meta-analysis on the relationship between breastfeeding and acute otitis media (AOM)	31
Table 2. Summary table for cohort studies on the relationship between breastfeeding and acute otitis media	32
Table 3. Summary of systematic review/meta-analysis on the relationship between breastfeeding and the risk of developing atopic dermatitis	36
Table 4. Summary of systematic review/meta-analysis on the relationship between breastfeeding and gastrointestinal (GI) infection.....	39
Table 5. Summary of systematic review/meta-analysis on the relationship between breastfeeding and lower respiratory tract disease	41
Table 6. Summary of systematic review/meta-analysis on the relationship between breastfeeding and asthma.....	47
Table 7. Summary of prospective cohort studies on the relationship between breastfeeding and asthma.....	48
Table 8. Summary of systematic reviews/meta-analyses on the relationship of breastfeeding and cognitive development	54
Table 9. Summary of studies on the relationship of breast milk feeding and cognitive development.....	56
Table 10. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and overweight or obesity	66
Table 11. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and total cholesterol and low-density lipoprotein cholesterol.....	69
Table 12. Summary table for systematic reviews/meta-analyses on the relationship between breastfeeding in infancy and blood pressure levels later in life.....	71
Table 13. Summary of systematic review/meta-analysis on the relationship between breastfeeding and long-term cardiovascular disease mortality	74
Table 14. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and type 1 diabetes	79
Table 15. Summary of case-control studies on the relationship between breastfeeding and type 1 diabetes.....	80
Table 16. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and type 2 diabetes	83
Table 17. Combined SES-adjusted ORs of ALL for the three case-control studies rated as good and fair methodological quality in the systematic review by Guise et al. (2005).....	87
Table 18. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and childhood leukemia.....	88
Table 19. Summary of case-control study on the relationship between breastfeeding and infant mortality.....	92
Table 20. Summary of systematic review/meta-analysis on the relationship between breastfeeding and Sudden Infant Death Syndrome (SIDS)	96

Table 21. Summary of case-control studies on the relationship between breastfeeding and sudden infant death syndrome (SIDS) included in the meta-analysis	97
Table 22. Meta-Analysis of four RCTs on the effects of breast milk feeding and NEC in preterm infants: Random Effects Model (D&L).....	99
Table 23. Summary of systematic review/meta-analyse on the relationship between breast milk feeding and necrotizing enterocolitis (NEC) in preterm infants.....	101
Table 24. Summary of studies on the relationship between breast milk feeding and necrotizing enterocolitis (NEC) in preterm infants.....	102
Table 25. Summary of primary studies on the relationship of breast milk feeding and cognitive development in preterm infants.....	106
Table 26. Summary of prospective cohort studies on the relationship between breastfeeding and returning to pre-pregnancy weights	114
Table 27. Summary of prospective cohort studies on the relationship between breastfeeding and postpartum weight/BMI changes	116
Table 28. Summary of systematic review/meta-analyse on the relationship between breastfeeding and maternal type 2 diabetes	122
Table 29. Summary of studies on the relationship between breastfeeding and maternal type 2 diabetes (NIDDM)	123
Table 30. Summary of case-control studies on the relationship between breastfeeding and risk of fracture in post-menopausal women.....	127
Table 31. Summary of prospective cohort studies on the relationship between duration of breastfeeding and the long-term changes of bone mineral density (BMD) or bone mineral content (BMC)	129
Table 32. Summary of the studies on the relationship between breastfeeding and postpartum depression	132
Table 33. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and the risk of breast cancer	139
Table 34. Summary of case-control studies on the relationship between breastfeeding and breast cancer.....	140
Table 35. Summary of case-control studies on the relationship between breastfeeding and maternal ovarian cancer	147
Table 36. Summary of case-control studies on the relationship between breastfeeding and maternal ovarian cancer by menopausal status.....	151
Table 37. Summary of case-control studies on the relationship between breastfeeding and maternal ovarian cancer by histologic type	152
Table 38. Summary of pooled-analysis of case-control studies on the relationship between breastfeeding and ovarian cancer by population subgroups	154

Appendixes

- Appendix A. MEDLINE® Search Strategy
- Appendix B. Sample data abstraction forms
- Appendix C. Evidence Tables
- Appendix D. List of Excluded Studies
- Appendix E. Peer Reviewers

Appendixes and Evidence Tables are provided electronically at
<http://www.ahrq.gov/downloads/pub/evidence/pdf/brfout/brfout.pdf>

Executive Summary

Introduction

The purpose of this report is to summarize the literature concerning the relationship of breastfeeding and various infant and maternal health outcomes. This report was requested by the Department of Health and Human Services (DHHS) Office on Women's Health and was conducted through the Evidence-based Practice Center (EPC) program at the Agency for Healthcare Research and Quality (AHRQ).

Methods

Two key questions are addressed:

1. What are the benefits and harms for infants and children in terms of short-term outcomes, such as infectious diseases (including otitis media, diarrhea, and lower respiratory tract infections), sudden infant death syndrome (SIDS) and infant mortality, and longer-term outcomes such as cognitive development, childhood cancer (including leukemia), type I and II diabetes, asthma, atopic dermatitis, cardiovascular disease (including hypertension), hyperlipidemia, and obesity, compared among those who mostly breastfeed, mostly formula feed, and mixed feed; and how are these outcomes associated with duration of the type of feeding? Do the harms and benefits differ for any specific subpopulations based on socio-demographic factors?
2. What are the benefits and harms on maternal health short-term outcomes, such as postpartum depression and return to pre-pregnancy weight, and long-term outcomes, such as breast cancer, ovarian cancer, diabetes and osteoporosis, compared among breastfeeding, formula feeding, and mixed feeding, and how are these associated with duration of the type of feeding? Do the harms and benefits differ for any specific subpopulations based on socio-demographic factors?

Approach to Evaluating the Literature

Inclusion Criteria. As it was not feasible to review the large number of primary studies that are relevant to all the outcomes of interest, we consulted the Office on Women's Health and the technical expert panel (TEP) and developed an approach that capitalized on the existing large number of systematic reviews/meta-analyses. For outcomes of interest that have previously been reviewed systematically, we have summarized the findings from those reviews. For acute otitis media, childhood asthma, cognitive development, SIDS, infant mortality, NEC, maternal breast cancer, return to pre-pregnancy weight, and maternal type 2 diabetes, we have also updated those systematic reviews with data from primary studies published subsequent to those reviews. For outcomes that have not been previously evaluated systematically (osteoporosis, ovarian cancer, postpartum depression, infant mortality), we have reviewed those primary studies that met our inclusion criteria. Studies that examined only formula-fed infants were excluded.

Definitions of Breastfeeding. The majority of the studies did not distinguish between exclusive and partially breastfed infants, or explain the difference between “breastfeeding” and “feeding of expressed breast milk.” We elected to use the term “breastfeeding” for studies in full-term infants and the term “human milk feeding” for studies in preterm infants. We elected to accept all definitions of “exclusive breastfeeding” as provided by the different study authors but qualified our conclusions with respect to those specific definitions.

Literature Search Strategy. Comprehensive literature searches of MEDLINE[®], CINAHL, and the Cochrane Database of Systemic Reviews took place in November of 2005. Search terms included subject headings and text words relevant to breastfeeding and the different outcomes. Supplemental searches on selected outcomes were conducted through May of 2006. Other relevant studies were identified by technical experts or in bibliographies of selected reviews.

Specific Inclusion Criteria for Health Conditions Evaluated. We included systematic reviews, meta-analyses, observational studies, randomized controlled trials, and comparative studies that evaluated the effects or associations of breastfeeding on outcomes of interest. All studies must have either a comparator arm that evaluated formula feeding or a comparator arm that evaluated different durations of breastfeeding. Only studies conducted in developed countries were used in updates of systematic reviews/meta-analyses and de novo reviews of primary studies.

Reporting of Evidence

Methodological Quality Grade of Individual Studies. We used a three category grading system (A, B, C) to denote methodological quality of each primary study. We did not evaluate the methodological quality of the individual studies in the systematic reviews/meta-analyses.

A (good): Least bias and results are valid; a primary study that adheres mostly to the commonly held concepts of high quality

B (fair/moderate): Susceptible to some bias, but not sufficient to invalidate the results; a primary study that does not meet all the criteria in category A

C (poor): Significant biases that may invalidate the results; a primary study with serious errors in design, analysis or reporting

Methodological Quality Grade of Systematic reviews/Meta-analyses. We used a similar scheme as above for grading systematic reviews/meta-analyses. But we supplemented the scheme with the MOOSE guideline (standards for reporting for meta-analysis in observational studies in epidemiology) and an additional checklist of items that we devised to evaluate the quality of the systematic review of observational studies. Items in this checklist included questions on the following: appropriate search strategy; justification for inclusion/exclusion criteria for studies; description of well-defined population, intervention/exposure, comparator, outcomes and study designs; effort to minimize errors in data extraction; assessment of quality of individual studies; consideration on the effect of confounders; combinability of the data for meta-analysis; assessment of statistical and clinical heterogeneity; reporting accuracies; and appropriateness of the conclusions based on the reported data.

Results

We screened over 9,000 abstracts. Forty-three primary studies on infant health outcomes, 43 primary studies on maternal health outcomes, and 29 systematic reviews or meta-analyses that covered approximately 400 individual studies were included in this review.

The association studies of breastfeeding and health outcomes mostly presented results as odds ratios. To facilitate interpretation of the odds ratio, we chose to present these data as a reduction in relative risk, estimated as “(1 – odds ratio) x 100%,” along with the corresponding 95% confidence interval (CI).

Full term Infant Outcomes

Acute Otitis Media. Our meta-analysis of five cohort studies of good and moderate methodological quality showed that breastfeeding was associated with a significant reduction in the risk of acute otitis media. Comparing ever breastfeeding with exclusive formula feeding, the risk reduction of acute otitis media was 23 percent (95% CI 9% to 36%). When comparing exclusive breastfeeding with exclusive formula feeding, either for more than 3 or 6 months duration, the reduction was 50 percent (95% CI 30% to 64%). These results were adjusted for potential confounders.

Atopic Dermatitis. One good quality meta-analysis of 18 prospective cohort studies on full term infants reported a reduction in the risk of atopic dermatitis by 42 percent (95% CI 8% to 59%) in children with a family history of atopy and exclusively breastfed for at least 3 months compared with those who were breastfed for less than 3 months. The meta-analysis did not distinguish between atopic dermatitis of infancy (under 2 years of age) and persistent or new atopic dermatitis at older ages. It has been postulated that the diagnosis of atopic dermatitis in patients younger than 2 years of age could be attributed to infectious etiologies, which may be prevented by breastfeeding. However, a stratified analysis by duration of followup found the risk reduction from breastfeeding was similar in subjects with less than 2 years compared with more than 2 years of followup.

Gastrointestinal Infections. For non-specific gastroenteritis, one systematic review identified three primary studies that controlled for potential confounders. These studies reported that there was a reduction in the risk of non-specific gastrointestinal infections during the first year of life in breastfed infants from developed countries. But a summary adjusted estimate taking into account potential confounders could not be determined because the studies did not provide usable quantitative data. However, a recent case-control study from England that took into account the role of potential confounders reported that infants who were breastfeeding had a 64 percent (95% CI 26% to 82%) reduction in the risk of non-specific gastroenteritis compared with infants who were not breastfeeding.

Lower Respiratory Tract Diseases. The summary estimate from a good quality meta-analysis of seven studies reported an overall 72 percent (95% CI 46% to 86%) reduction in the risk of hospitalization due to lower respiratory tract diseases in infants less than 1 year of age who were exclusively breastfed for 4 months or more. The results remained consistent after adjustment for potential confounders.

Asthma. The studies on asthma were equivocal. A previously published good quality meta-analysis reported a moderate protective effect and four recent primary studies reaching mixed conclusions, including two studies finding an increased risk of asthma associated with breastfeeding. We updated the meta-analysis with the new studies. Our analysis showed that breastfeeding for at least 3 months was associated with a 27 percent (95% CI 8% to 41%) reduction in the risk of asthma

in those subjects without a family history of asthma compared with those who were not breastfed. For those with a family history of asthma, there was a 40 percent (95% CI 18% to 57%) reduction in the risk of asthma in children less than 10 years of age who were breastfed for at least 3 months compared with those who were not breastfed. However, the relationship between breastfeeding and the risk of asthma in older children and adolescents remains unclear and will need further investigation.

Cognitive Development. One well-performed sibling analysis and three prospective cohort studies of full-term infants, all conducted in developed countries, adjusted their analyses specifically for maternal intelligence. The studies found little or no evidence for an association between breastfeeding in infancy and cognitive performance in childhood. Most of the published studies adjusted their analyses for socioeconomic status and maternal education but not specifically for maternal intelligence. For those studies that reported a significant effect after specific adjustment for maternal intelligence, residual confounding from other factors such as different home environments cannot be ruled out.

Obesity. Three meta-analyses of good and moderate methodological quality reported an association of breastfeeding and a reduction in the risk of obesity in adolescence and adult life compared with those who were not breastfed. One study reported the reduction in the risk of overweight/obesity in breastfeeders compared with non-breastfeeders was 24 percent (95% CI 14% to 33%); another study reported 7 percent (95% CI 1% to 12%). Both of these estimates took into account the role of potential confounders. Furthermore, they also showed that the magnitude of association decreased when more confounders were entered into the analyses. The third study used meta-regression and found a 4 percent reduction in the risk of being overweight in adult life for each additional month of breastfeeding in infancy. Overall, there is an association between a history of breastfeeding and a reduction in the risk of being overweight or obese in adolescence and adult life. One should be cautious in interpreting all these associations because of the possibility of residual confounding.

Risk of Cardiovascular Diseases. Results from two moderate quality meta-analyses concluded that there was a small reduction of less than 1.5 mm Hg in systolic blood pressures and no more than 0.5 mm Hg in diastolic blood pressures among adults who were breastfed in their infancy compared with those who were formula-fed. The association weakened after stratification by study size, suggesting the possibility of bias in the smaller studies.

One meta-analysis of cohort and case-control studies reported that there was a reduction in total and LDL cholesterol levels by 7.0 mg/dL and 7.7 mg/dL, respectively, in adults who were breastfed during infancy compared with those who were not. However, these findings were based on data from adults with a wide age range. The analysis did not segregate the data according to gender and potential confounders were not explicitly analyzed. Detailed information (e.g., fasting or non-fasting) on the collection of specimen for cholesterol testing was not included. Because of these deficiencies, the correct characterization of a relationship between breastfeeding and adult cholesterol levels cannot be determined at this time.

One meta-analysis found little or no difference in all-cause and cardiovascular mortality between adults who were breastfed during infancy and those who were not. There were possible biases and limitations in the studies reviewed, however. Presence of statistical heterogeneity across studies suggests that it may not have been appropriate to combine estimates from individual studies into one summary estimate. Because of these reasons, no definitive conclusion could be drawn regarding the relationship between a history of breastfeeding and cardiovascular mortality.

In summary, the relationship between breastfeeding in infancy and the risk of cardiovascular diseases cannot be confidently characterized at this time and will need further investigation.

Type 1 Diabetes. Two moderate quality meta-analyses suggest that breastfeeding for at least 3 months reduced the risk of childhood type 1 diabetes compared with breastfeeding for less than 3 months. One reported a 19 percent (95% CI 11% to 26%) reduction; the other reported a 27 percent (95% CI 18% to 35%) reduction. In addition, findings from five of six studies published since the meta-analyses reported similar results. However, these results must be interpreted with caution because of the possibility of recall biases and suboptimal adjustments for potential confounders in the studies.

Type 2 Diabetes. In one well-performed meta-analysis of seven studies of various designs, breastfeeding in infancy was associated with a 39 percent (95% CI 15% to 56%) reduced risk of type 2 diabetes in later life compared with those who were not. However, only three of seven studies adjusted for all the important confounders such as birth weight, parental diabetes, socioeconomic status, and individual or maternal body size. Though the crude and adjusted estimates did not differ in these three studies, the lack of adjustments for potential confounders such as birth weight and maternal factors by all studies could exaggerate the magnitude of an association.

Childhood Leukemia. The published studies on childhood acute lymphocytic leukemia (ALL) were equivocal; a good quality meta-analysis reported a moderate protective effect from breastfeeding and the other good quality systematic review reached the opposite conclusion. We conducted a meta-analysis including only good and fair quality case-control studies identified in the systematic review, since the meta-analysis did not provide methodological quality grading of primary studies. We found breastfeeding of at least 6 months duration was associated with a 19 percent (95% CI 9% to 29%) reduction in the risk of childhood ALL. The previous meta-analysis also reported an association between breastfeeding of at least 6 months duration and a 15 percent reduction (95% CI 2% to 27%) in the risk of acute myelogenous leukemia (AML). Overall there is an association between a history of breastfeeding for at least 6 months duration and a reduction in the risk of both leukemias (ALL and AML).

Infant Mortality. One study of moderate methodological quality evaluated the relationship between breastfeeding and infant mortality. The study reported a protective effect of breastfeeding in reducing infant mortality after controlling for some of the potential confounders. However, in subgroup analyses of the study, the only statistically significant association reported was between “never breastfed” and Sudden Infant Death Syndrome (SIDS) or the risk of injury-related deaths. Because of the limited data in this area, the relationship between breastfeeding and infant mortality in developed countries remains unclear. Further investigation is needed.

Sudden Infant Death Syndrome (SIDS). We conducted a meta-analysis by including only studies that reported clear definitions of exposure, outcomes, and results adjusted for well-known confounders or risk factors for SIDS. Our meta-analysis of seven case-control studies found that a history of breastfeeding was associated with a 36 percent (95% CI 19% to 49%) reduction in the risk of SIDS compared to those without a history of breastfeeding.

Preterm Infant Outcomes

Cognitive Development. No definitive conclusion can be made regarding the relationship between breast milk feeding and cognitive development in preterm infants. One meta-analysis reported a five points advantage in standardized mean score and one systematic review identified one primary study that reported an eight points advantage in IQ in preterm or low birth weight infants who received breast milk feeding. In three of four primary studies of moderate quality that controlled for either maternal education or maternal intelligence, the advantage from breastfeeding was reduced to a statistically non-significant level after adjustment. The roles of maternal intelligence and home

environment should be accounted for in future studies on breastfeeding and cognitive development. Keeping in mind that cognitive function measured at an early age is not necessarily predictive of later cognitive ability, one should also consider carefully the timing and the selection of appropriate testing instrument in future studies.

Necrotizing Enterocolitis (NEC). Our meta-analysis of four randomized controlled trials of breast milk versus formula in comparing the outcome of NEC demonstrated that there was a marginally statistically significant association between a history of breast milk feeding and a reduction in the risk of NEC ($P = 0.04$). The estimate of the reduction in relative risk ranged from 4 percent to 82 percent. The absolute risk difference between the two groups was 5 percent. Because of the high case-fatality rate of NEC, this difference is a meaningful clinical outcome. The wide range of the estimate reflects the relatively small number of total subjects in the studies and the small number of events. One must also be cognizant of the heterogeneity underlying these trials in interpreting the findings of the meta-analysis. Examples of which included gestational age that ranged from 23 to more than 33 weeks; birth weight ranged from less than 1,000 g to more than 1,600 g; and some trials included only “healthy” infants, while others included both “healthy” and “ill” infants.

Maternal Outcomes

Return to Pre-pregnancy Weight. Three moderate quality prospective cohort studies reported less than 1 kg weight change from pre-pregnancy or first trimester to 1 to 2 year postpartum period in mothers who breastfed. Results from four moderate quality prospective cohort studies showed that the effects of breastfeeding on postpartum weight loss were unclear. Results from all seven studies consistently showed that many factors other than breastfeeding had larger effects on weight retention or postpartum weight loss. Methodological challenges in these studies included the accurate measurement of weight change, adequate control for numerous covariables including the amount of pregnancy weight gain, and quantifying accurately the exclusivity and the duration of breastfeeding.

Maternal Type 2 Diabetes. Two large cohorts from a high quality longitudinal study of 150,000 parous women in the United States examined the relationship between breastfeeding and the risk of maternal type 2 diabetes. In parous women without a history of gestational diabetes, each additional year of breastfeeding was associated with a 4 percent (95% CI 1% to 9%) reduced risk of developing type 2 diabetes in the first cohort and a 12 percent (95% CI 6% to 18%) reduced risk in the second cohort. In women with a history of gestational diabetes, breastfeeding had no significant effect on the already increased risk of diabetes. Because only nurses were included in the cohorts, generalization of findings to the rest of the population must be done with care.

Osteoporosis. There is little or no evidence from six moderate quality case-control studies for an association between lifetime breastfeeding duration and the risk of fractures due to osteoporosis. In two of three moderate or good quality prospective cohort studies using bone mineral density as a surrogate for osteoporosis, lactation does not appear to have an effect on long-term changes in bone mineral densities. The third study found a small decrease in the bone mineral contents in the distal radius with increased duration of breastfeeding, but no significant changes in bone mineral contents in the femoral neck or the trochanter.

Postpartum Depression. Four prospective cohort studies of moderate methodological quality reported on the relationship between a history of breastfeeding and postpartum depression. None of the studies explicitly screened for depression at baseline before the initiation of breastfeeding and none of them provided detailed data on breastfeeding. Three of the four studies found an association between a history of short duration of breastfeeding or not breastfeeding with postpartum depression.

The results were adjusted for socio-demographic and obstetric variables. More investigation will be needed to determine the nature of this association. It is plausible that postpartum depression led to early cessation of breastfeeding, as opposed to breastfeeding altering the risk of depression. Both effects might occur concurrently.

Breast Cancer. Two meta-analyses of moderate methodological quality concluded that there was a reduction of breast cancer risk in women who breastfed their infants. The reduction in breast cancer risk was 4.3 percent for each year of breastfeeding in one meta-analysis and 28 percent for 12 or more months of breastfeeding in the other. In addition, one of the two meta-analyses and another systematic review reported decreased risk of breast cancer primarily in premenopausal women. Findings from primary studies published after the meta-analyses concurred with the findings from the earlier meta-analyses. In summary, consistent evidence from these studies suggests that there is an association between breastfeeding and a reduced risk of breast cancer.

Ovarian Cancer. We reviewed 15 case-control studies that examined the relationship between breastfeeding and the risk of ovarian cancer, and performed meta-analyses in nine studies that adjusted for potential confounders. The overall result from the nine studies showed an association between breastfeeding and a 21 percent (95% CI 9% to 32%) reduction in the risk of ovarian cancer, compared to never breastfeeding. Because not all the studies reported similar comparisons of breastfeeding durations, we had to estimate the comparable risks in five studies. Excluding these five studies from the meta-analysis results in loss of statistical significance for this association.

There was indirect evidence for a dose-response relationship between breastfeeding and a reduced risk of ovarian cancer. Breastfeeding of more than 12 months (cumulative duration) was associated with a reduced risk of ovarian cancer, compared to never breastfeeding. The 12-month cutoff was arbitrary, and the odds ratios were estimated in half of these studies.

Overall, there is evidence to suggest an association between breastfeeding and a reduction in the risk of maternal ovarian cancer. Because of the aforementioned limitations, one must be cautious in interpreting this association.

Discussion

Limitations

With the availability of many published systematic reviews on breastfeeding, we used this literature as the evidence for a large number of outcomes, supplemented by updates of these systematic reviews with new primary studies. Even though we have assessed the reporting quality of these systematic reviews (using standards of reporting of systematic reviews of observational studies (MOOSE statement), and additional parameters that we devised), we cannot reliably know the validity of the reported summary data without knowing the details of the primary studies. It should also be stressed that a well-performed systematic review does not necessarily imply that the body of evidence for a particular outcome of interest is of high quality. Any systematic review is limited by the quality of the primary studies included in the review. Unless the method used to assess the quality of the primary studies is transparent and the details made available for examination, it would be difficult to reliably determine the validity of the conclusions.

The breastfeeding literature is primarily comprised of observational studies, either cohort or case-control studies. There are a number of potential deficiencies related to the observational study designs that could limit the internal validity and the generalizability of the findings. Some of these potential deficiencies include (1) misclassification of exposure; (2) confounding from the process of self-selection; and (3) residual confounding.

We have summarized the effects of breastfeeding (or breast milk feeding) on a large number of infant and maternal outcomes. Some of the outcomes are well defined and specific (e.g., childhood acute lymphocytic leukemia, breast cancer); and some are not so well defined and non-specific (e.g., asthma, non-specific gastrointestinal infections). When the reported outcome is well defined and specific, it lends confidence that the effect reported is valid for that outcome. When the reported outcome is not well defined, one might have some reservation regarding the validity of the measured effect for that outcome. For all the above reasons, we find that there is a wide range of quality of evidence for the different outcomes examined in this review.

An important area of research that is not systematically reviewed in this report is the use of breastfeeding promotion intervention trial to measure health effects (this topic is not part of the scope of this report and it will be covered in a separate report). The best known of these types of studies is the Promotion of Breastfeeding Intervention Trial (PROBIT) conducted in the Republic of Belarus. Data from this study provided good evidence that breastfeeding is associated with a reduction in the risk of gastrointestinal infection and atopic dermatitis.

Lastly, the outcomes analyzed in this review represent only a portion of all possible health outcomes related to breastfeeding reported by investigators worldwide. To work within the constraints of resources, we relied on the advice from our panel of technical experts in finalizing the list of outcomes included in this review. Thus, some important outcomes (e.g., growth and nutrition) have, by necessity, not been included in this review.

Future Research

Observational studies will remain the major source of information in this field. Clear subject selection criteria, adopting a common definition of “exclusive breastfeeding”, reliable collection of feeding data, specific and properly quantifiable outcomes of interest, controlling for important potential confounders including child-specific factors, and blinded assessment of the outcome measures will help immeasurably to improve the quality of these studies.

Sibling analysis provides a method to control for hereditary and household factors that are important in certain outcomes, provided that those factors are similar for the siblings of interest. Although such analysis may be less susceptible to confounders and effect modifiers that are shared by siblings, one must remember that it is not immune to biases. This method should be used when the appropriate data are available.

Cluster randomized controlled studies similar to the Belarus trial will provide understanding of the effectiveness of various breastfeeding promotion interventions. Any substantial differences in the degree of breastfeeding between the two groups as a result of the intervention will provide further opportunity to investigate any disparity in health outcomes between the two groups.

Evidence Report

Chapter 1. Introduction

The Department of Health and Human Services (DHHS) Office on Women's Health has requested an evidence report from the Agency for Healthcare Research and Quality (AHRQ) through the Evidence-based Practice Center program (EPC) that would critically examine the literature concerning the relationship of breastfeeding and various infant and maternal health outcomes. EPC evidence reports summarize evidence addressing specific key questions; these reports do not make clinical practice or health policy recommendations.

Breast milk is the natural nutrition for all infants. According to the American Academy of Pediatrics (AAP), it is the preferred choice of feeding for all infants.¹ The goals of *Healthy People 2010* for breastfeeding are an initiation rate of 75 percent and continuation of breastfeeding of 50 percent at 6 months and 25 percent at 12 months postpartum.² National Immunization Survey of U.S. children in 2005 (NIS 2005) indicated that 73 percent had ever been breastfed. The percentage of infants who continued to breastfeed to some extent is 39 percent at 6 months and 20 percent at 12 months (www.cdc.gov/breastfeeding/data/NIS_data/data_2005.htm).

In addition to providing essential nutrients to infants, benefits of breastfeeding for both children and their mothers have been reported. Reports of the benefits for children include decreases in incidence of otitis media and gastroenteritis,³ lower risk of obesity,^{4,5} and lower risk of asthma.⁶ Other benefits reported include decreased rates of sudden infant death syndrome, reduction in the incidence of type 1 and type 2 diabetes mellitus, certain types of cancer, and improved performance on certain tests of cognitive development.⁷

Reported benefits for mothers who breastfed their infants include increased postpartum uterine activity (inferentially this would lead to reduced postpartum blood loss),⁸ greater weight loss postpartum compared with mothers who bottle-fed their infants,⁹ decreased incidence of premenopausal breast cancer¹⁰ and decreased incidence of ovarian cancer.¹¹

In 2000, the DHHS Office on Women's Health, in cooperation with the Surgeon General of the United States and several governmental and non-governmental agencies, published the first departmental policy on breastfeeding, the *HHS Blueprint for Action on Breastfeeding* (www.womenshealth.gov). The DHHS Office on Women's Health endorses the recommendation from the American Academy of Pediatrics (AAP), American Academy of Family Physicians (AAFP), American College of Obstetricians and Gynecologists (ACOG), Association of Women's Health, Obstetrical, and Neonatal Nurses (AWHONN), Le Leche League International, National Medical Association (NMA), and many other health organizations, that mothers exclusively breastfeed for 6 months. The DHHS Office on Women's Health has commissioned a review to systematically examine the evidence for the effects of breastfeeding. The DHHS Office on Women's Health has also requested that the focus of the review be on studies from developed countries (i.e., "high income" classification by World Bank)* as the findings from those studies are deemed more directly applicable to population in this country.

As it is unethical to randomize subjects into breastfeeding versus non-breastfeeding groups (although there were some randomized controlled trials (RCTs) in the 1980s on preterm infants whose mothers desired not to breastfeed, their infants were randomized into those who received

* http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20421402~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html#High_income

donor breast milk versus those who received preterm formula^{12,13}), much of the evidence on the benefits of breastfeeding came from observational studies. Observational studies are subject to confounding. One of the well-known confounders in breastfeeding research is demographic difference between mothers who breastfeed and those who chose not to breastfeed due to self-selection. Consistent with previously reported data,¹⁴ NIS 2005 showed that mothers who breastfeed tend to be white (versus non-Hispanic black or African American), older, more educated, and in a higher socioeconomic stratum (cdc.gov/breastfeeding/data/NIS_data/data_2005.htm). While it is possible to control for some of these demographic factors, it is not possible to control for behavioral or attitudinal factors intrinsic in the desire to breastfeed. Some authors have proposed strict standards in evaluating the quality of observational studies. These standards should include the quality of the feeding data, a clear definition of the outcome, the elimination of systematic differences in outcome assessment between comparison groups (detection bias), and the control of potential and well-known confounders. The feeding data should clearly define whether it was prospectively or retrospectively collected, whether there was a precise definition of exclusive breastfeeding, and whether the duration of breastfeeding was reported.^{15,16}

Large number of infant and maternal outcomes has been examined in relation to a history of breastfeeding. It was not feasible to review all possible outcomes in mothers and children for this report; we sought guidance from our panel of technical experts in the field of breastfeeding research in deciding on the specific outcomes to review. After taking into consideration the following factors: relevance and importance of outcome in a developed country, date (recent or old) of the last systematic review on the outcome, availability or non-availability of data from a developed country, consistency or inconsistency of outcomes in previously reported studies, and consideration of the possibility that breastfeeding may have potential harms as well as benefits, the following outcomes from developed countries have been designated for review: for term infants, infectious diseases (including otitis media, diarrhea, and lower respiratory tract infections), sudden infant death syndrome, infant mortality, cognitive development, childhood cancer (including leukemia), type 1 and 2 diabetes, asthma, atopic dermatitis, cardiovascular disease (including hypertension), hyperlipidemia, and obesity; for preterm infants, necrotizing enterocolitis (NEC) and cognitive development; for mothers, post-partum depression, return to pre-pregnancy weight, breast cancer, ovarian cancer, diabetes and osteoporosis.

It is also outside the scope of this report to examine the biological mechanisms underpinning the effects of breast milk and therefore, studies on individual components of breast milk will not be part of this report. Lastly, studies on the effectiveness of interventions to promote and support breastfeeding are not systematically covered in this review, as this topic will be reviewed for a subsequent report. However, there is good quality evidence to support that some of these interventions do lead to an increase in breastfeeding rates and also an improvement of certain health outcomes in the study populations. Details of one landmark study¹⁷ and its implications for future research in the study of the effects of breastfeeding will be discussed in some details in this report.

Chapter 2. Methods

Overview

This evidence report on breastfeeding and health outcomes in infants and mothers is based on a systematic review of the literature. To identify the specific issues central to this report, the Tufts-New England Medical Center (Tufts-NEMC) Evidence-based Practice Center (EPC) held teleconferences with a panel of technical experts (TEP) and various stakeholders. A comprehensive search of the medical literature was conducted to identify studies addressing the key questions. Evidence tables of study characteristics and results were compiled, and the methodological quality of the studies was appraised. Study results were summarized with qualitative reviews of the evidence, summary tables, and quantitative summary data, when appropriate.

A number of individuals and groups supported the Tufts-NEMC EPC in preparing this report. The TEP served as our science partner. Technical experts and representatives from the Agency for Healthcare Research and Quality (AHRQ), DHHS Office on Women's Health, The National Institute of Child Health and Human Development (NICHD), Centers for Disease Control and Prevention (CDC), American Academy of Pediatrics (AAP), American College of Obstetrics and Gynecology (ACOG), Association of Women's Health, Obstetrics and Neonatal Nurses (AWHONN), and La Leche League worked with the EPC staff to refine key questions, identify important issues, and define parameters for literature review in this report.

In the early phase of exploring the literature available for this report, it was soon discovered that there was a large number of primary studies and systematic reviews/meta-analyses on the various outcomes of interest. As it was not feasible to review all the primary studies addressing the outcomes of interest, therefore, in consultation with the Office on Women's Health and the TEP, we developed an approach that capitalized on the existing systematic reviews/meta-analyses. For outcomes of interest that had previously been reviewed systematically, we assessed the quality of those reviews and summarized their findings. For selected infant (necrotizing enterocolitis, cognitive development, acute otitis media, asthma, type 1 and 2 diabetes, SIDS) and maternal (weight changes, type 2 diabetes, breast cancer) outcomes, in addition to reporting on the existing systematic reviews, we also updated them by summarizing the relevant primary studies that were published after those reviews. For outcomes of interest (osteoporosis, ovarian cancer, postpartum depression, infant mortality) that had not been reviewed systematically, we reviewed all the relevant primary studies that met our inclusion criteria.

Key Questions Addressed in This Report

Two key questions are addressed in this report. Question 1 pertains to infant outcomes and question 2 pertains to maternal outcomes. The key questions are:

1. What are the benefits and harms for infants and children in terms of short-term outcomes, such as infectious diseases (including otitis media, diarrhea, and lower respiratory tract infections), sudden infant death syndrome and infant mortality, and longer-term outcomes such as cognitive development, childhood cancer (including leukemia), type 1 and 2 diabetes, asthma, atopic dermatitis, cardiovascular disease (including hypertension), hyperlipidemia, and obesity, compared among those who mostly breastfeed, mostly

formula feed, and mixed feed; and how are these outcomes associated with duration of the type of feeding? Do the harms and benefits differ for any specific subpopulations based on socio-demographic factors?

2. What are the benefits and harms on maternal health short-term outcomes, such as post-partum depression and return to pre-pregnancy weight, and long-term outcomes, such as breast cancer, ovarian cancer, type 2 diabetes mellitus and osteoporosis, compared among breastfeeding, formula feeding, and mixed feeding, and how are these associated with duration of the type of feeding? Do the harms and benefits differ for any specific subpopulations based on socio-demographic factors?

It should be emphasized that the focus of this review is on the effects of breast milk feeding, not formula feeding. However, many studies did not distinguish between exclusive and partially breastfed infants; presumably, some of the effects reported from observational studies were from infants who received both breast milk and formula milk feedings. Studies that examined only formula fed infants were not included in this report. Lastly, studies on infant and maternal health outcomes of interventions to promote and support breastfeeding were not systematically covered in this review as that subject will be covered in a separate report. However, our panel of technical experts felt that the study of breastfeeding promotion in infants from Belarus¹⁷ was a landmark study and offered new directions into research on effects from breast milk that it warrants discussion in this report. The details of that study are described in the section Other Research in the results chapter.

Definitions of Breastfeeding in This Report

None of the studies in this review explicitly examined the difference between “breastfeeding” an infant (infant suckling at her/his mother’s nipple) and “feeding of expressed breast milk” to an infant. To distinguish between the two forms of feedings, we elected to use the term “breastfeeding” when the studies concerned primarily full-term infants (presumably they were, indeed, breastfed) and the term “breast milk feeding” when the studies concerned primarily preterm infants (as most of them received breast milk initially either by gavage- or by bottle-feeding). For term infants, “bottle-feeding” is used synonymously with “formula feeding.”

Definitions of “exclusive breastfeeding” varied widely in the literature. They ranged from “no supplement of any kind including water while breastfeeding” to “occasional formula is permissible while breastfeeding.” We elected to accept all definitions of “exclusive breastfeeding” as provided by the different study authors, but we qualified our findings by the details regarding those definitions.

Literature Search Strategy

We conducted a comprehensive literature search to address the two key questions. The EPC used the Ovid search engine to conduct searches on the MEDLINE[®] database, CINAHL database, and the Cochrane Database of Systemic Reviews. A wide variety of search terms were used to capture the many potential sources of information related to the myriad of different outcomes (see Appendix A).^{*} But the different outcomes were always searched in conjunction

^{*} Appendixes and Evidence Tables cited in this report are provided electronically at <http://www.ahrq.gov/clinic/tp/brfouttp.htm>

with the following: “breastfeeding,” “breast milk feeding,” “breast milk,” “human milk,” “nursing”, and “lactation”. Literature search of the outcomes alone without references to breast milk feeding was not conducted. The search included citations from 1966 to November of 2005. Updated searches on selected outcomes took place in April and May of 2006. We also supplemented our computer search by examining the bibliographies of the review articles. We also included articles suggested by reviewers, provided that the articles met the inclusion criteria for this review. For outcomes that were not slated for updates, additional articles suggested by reviewers were also included as an addendum if they provided useful information. We did not make efforts to identify unpublished studies.

Study Selection

Selection of Outcomes for This Review

The TEP offered advice on selection of outcomes for review. Final selection of the list of outcomes for review took into account the following factors: the importance of the outcome, whether a systematic review of the outcome has previously been reported from a developed country, whether the existing systematic review of the outcome is outdated, whether the relationship between breastfeeding and the outcome is thought to be equivocal, whether a large number of primary studies has been published recently on the outcome, and the total number of outcomes that could be adequately reviewed for this report given the time constraint.

We included the following outcomes in this report:

Term infant outcomes: acute otitis media, hospitalization for lower respiratory tract infection, gastrointestinal infection, hypertension, cardiovascular diseases, hyperlipidemia, asthma, atopic dermatitis, type 1 and 2 diabetes, obesity, sudden infant death syndrome (SIDS), infant mortality, cognitive development, and childhood cancer (including leukemia)

Preterm infant outcomes: necrotizing enterocolitis (NEC) and cognitive development

Maternal outcomes: maternal weight changes, breast cancer, ovarian cancer, post-partum depression, osteoporosis, and type 2 diabetes

Abstract Screening

All abstracts identified through the literature search were screened. At this stage, eligible studies included all English language primary experimental or observational studies that reported any health outcomes in human subjects in relation to a history of breast milk feeding. As the inclusion criteria were broad at this stage, the abstracts rejected at this stage did not undergo a second rescreening process. Abstracts that were accepted at this stage were examined a second time by different reviewers and categorized into the different outcomes of interest.

Full Article Inclusion/Exclusion Criteria

Articles that passed the abstract screening process were retrieved and the full articles were reviewed for eligibility. Full articles were examined only once unless the articles were equivocal for inclusion or exclusion. In that event, the article in question was screened again by a different reviewer and a consensus was reached after discussion with the first reviewer.

Because the outcomes selected ranged from very broad topic with common occurrence (e.g., non-specific gastrointestinal infection) to a narrowly focused topic with relatively few occurrences (e.g., SIDS), the types of studies available for each outcome varied widely in the distribution of study designs, sample sizes, and quality of breastfeeding data, it was not possible (nor desirable) to design a strict set of inclusion and exclusion criteria that would be applicable to all outcomes. Therefore, additional inclusion/exclusion criteria germane to the specific outcome were also described in the Results section under each outcome.

General inclusion criteria for the studies are as follow:

Study Design. Systematic reviews, experimental (randomized controlled trials) and observational studies (prospective cohort and case-control studies only)

Population. Healthy term infants in developed countries; preterm infants in developed countries (for NEC and cognitive development); healthy mothers in developed countries

Intervention/Exposure. Breastfeeding, breast milk feeding (maternal term and preterm milk, banked term and preterm milk, fortified or unfortified), exclusive or mixed feeding

Comparator. Formula feeding (preterm or term formula, fortified or unfortified)

Data Extraction and Analysis

For those outcomes that have been subjected to a systematic review/meta-analysis, we summarized their results into our report. In addition to the results from the systematic reviews, we have also extracted and summarized the relevant data from primary studies that were published after the latest search dates of the reports for the following infant and maternal outcomes: acute otitis media, childhood asthma, cognitive development, SIDS, infant mortality, NEC, and maternal breast cancer.

For systematic reviews/meta-analyses that reported data from both developed and developing countries, we reported only those results pertaining to developed countries, if the reported data permitted us to do so. In those instances where that were not possible, we noted that fact as a limitation of our findings.

For outcomes that have multiple systematic reviews, we noted the overlapping studies and examined whether their findings were interpreted similarly or differently across reviews and reported our analyses.

For NEC, maternal weight changes, and acute otitis media, in order to better clarify the overall findings, we also extracted relevant data from the primary studies cited in the systematic reviews and combined them with data from the primary studies that were published after the latest search dates of the reports and reanalyzed the data.

For the remaining included outcomes, we extracted and summarized the relevant data from the primary studies.

Data forms were developed separately for extraction of systematic reviews/meta-analysis and primary studies. For systematic reviews/meta-analysis, items extracted were: databases searched, study design, population characteristics, descriptions of intervention/exposure, models used for meta-analysis, results, and authors' conclusions. We reported the estimates in the meta-analyses. We also reported any attempt by the authors of the meta-analyses to explore heterogeneity using sub-group analyses or meta-regression. For primary studies, items extracted were: study design, population characteristics, eligibility criteria, descriptions of intervention/exposure, any adjustments for confounders, and results.

Meta-Analysis

We used meta-analysis to expand on the individual studies' findings, if it was appropriate and feasible to do so. Minimal criteria for meta-analysis are comparable groupings, similar study designs, and quantifiable outcome data. Secondary criteria for consideration of meta-analysis are similar study quality, similar statistical adjustment of outcomes, and other factors. Before combining the reported odds ratios or risk ratios reported in the individual studies from the previous meta-analysis with the estimates from the updated primary studies into a new summary odds ratio or risk ratio, we verified the previous reported odds ratios or risk ratios by examining the data from the original studies.

We used the DerSimonian and Laird's random effects model for all meta-analyses.¹⁸ The random effects model assigns a weight to each study based both on the individual study variance and the between-study heterogeneity. Compared with the fixed effect model, the random effects model is more conservative in that it generally results in broader confidence intervals when between-study heterogeneity is present. We tested for heterogeneity using Cochran's Q and assessed its extent with I^2 , which evaluates the proportion of between study variability that is attributed to heterogeneity rather than chance.^{19,20} Intercooled Stata 8.2 was used for the calculations and graphics.

Grading of Studies Analyzed in This Evidence Report

Studies accepted in evidence reports have been designed, conducted, analyzed, and reported with various degrees of methodological rigor and completeness. Deficiencies in any of these processes may lead to biased reporting or interpretation of the results. While it is desirable to have a simple evidence grading system using a single quantity, the quality of evidence is multi-dimensional. A single metric cannot adequately capture information needed to interpret a clinical study. However, grading of information can help the reader to interpret the studies properly.

Grading of Systematic Reviews/Meta-Analyses

We assessed the methodological quality of studies based on predefined criteria. For the assessment of systematic reviews, the criteria for methodological quality was based on the QUOROM guidelines for meta-analyses and systematic reviews of RCTs (a checklist organized into 21 headings and subheadings for the preferred way to present the abstract, introduction, methods, results, and discussion sections of a report of a meta-analysis),²¹ and reporting guidelines for meta-analysis in observational studies in epidemiology (MOOSE) (a checklist for the specifications for presenting background, search strategy, methods, results, discussion, conclusion, and assessment of quality of individual studies and bias (e.g., publication bias)^a).²² As the QUOROM and the MOOSE statements were primarily concerned with the reporting standards of the reviews, we have also supplemented those criteria with our own checklist of items designed to evaluate the quality of the systematic review of observational studies (see Appendix B^b for details). Items in this checklist consisted of questions on appropriate search strategy;

^a Publication bias refers to selective publication of studies according to results. When studies with non-significant or negative results are not published, the available studies may overestimate the effect.

^b Appendixes and Evidence Tables cited in this report are provided electronically at <http://www.ahrq.gov/clinic/tp/brfouttp.htm>.

justification for inclusion/exclusion criteria; how well-defined were the population, intervention, comparator, outcomes and study design; effort to minimize errors in data extraction; assessment of individual study quality; consideration on the effect of confounders; combinability of the data for meta-analysis; assessment of statistical and clinical heterogeneity; reporting accuracies; and appropriateness of the conclusions based on the reported data.

We applied a three category summary grading system (A, B, C) to each systematic review/meta-analysis:

A (good)

Category A studies have the least bias and results are considered valid. A study that adheres mostly to the commonly held concepts of high quality including the following: a rigorously conducted systematic review or meta-analysis; clear description of the population, setting, interventions and comparison groups; clear description of the content of the comparison groups; appropriate measurement of outcomes; appropriate statistical assumptions and analytic methods and reporting; appropriate consideration and adjustment for potential confounders; rigorous assessment of individual study quality; no reporting errors; and well-reasoned conclusions based on the data reported.

B (fair/moderate)

Category B studies are susceptible to some bias, but not sufficient to invalidate the results. They do not meet all the criteria in category A because they have some deficiencies, but none of which are likely to cause major biases. The study may have suboptimal adjustment for potential confounders. The study may also be missing information, making it difficult to assess limitations and potential problems.

C (poor)

Category C studies have significant biases that may invalidate the results. The study either did not consider potential confounders or did not adjust for them appropriately. These studies have serious errors in design, analysis or reporting; have large amounts of missing information, or discrepancies in reporting.

It should be noted that while we assessed the methodological quality of the systematic reviews or meta-analyses, it was not possible to evaluate the quality of the primary studies included in those reviews/analyses, as we did not examine those studies first hand. For systematic reviews/meta-analyses that had equivocal grading between moderate and poor, the results and the reasons for the initial grade assignments were presented to the entire group of project investigators and the final grades were adjudicated.

Grading of Individual Primary Studies in Updates and New Reviews

A well-performed RCT with proper randomization, allocation concealment, clear definitions of breastfeeding exposure compared with non-breastfeeding, and blinded assessment of outcomes will yield the best evidence in supporting the causality of breast milk in affecting health outcomes. But with the recognized benefits of breast milk, this approach is ethically not feasible. Other types of studies involve following the health outcomes from randomization of intervention to promote and support breastfeeding (e.g., Belarus study¹⁷), this type of study will yield indirect evidence for the relationship between breastfeeding and health

outcomes provided that there is a differential effect from the intervention on breastfeeding rates between the comparison groups. Prospective observational cohort studies with proper adjustment of potential confounders provide the bulk of data in this field. However, the possibility of residual confounding that could explain the observed association between breastfeeding and the specific health outcomes can never be completely ruled out. Case-controlled design is an even less attractive option because of the concern for case selection bias and suboptimal matching to control subjects.

For the assessment of RCTs, the criteria were based on the CONSORT statement for reporting RCTs (a checklist with specifications for reporting all aspects of a trial).^{23,24} We mainly considered the methods used for randomization, allocation concealment, and blinding as well as the use of intention-to-treat analysis, the report of well-described valid primary outcomes, and the dropout rate. For non-randomized trials, we used the report of eligibility criteria and assessed the adequacy of controlling for differences between comparative groups in terms of baseline characteristics and prognostic factors. We also considered the report of intention-to-treat analysis, and the crossovers when so designed, as well as important differential loss to followup between the comparative groups or overall high loss to followup. The validity and the adequate description of outcomes and results were also assessed. For the assessment of prospective cohorts and case-control studies (cross-over design and retrospective cohort studies were excluded from this review), we used a rating checklist largely based on the Newcastle-Ottawa Quality Assessment scales for cohort and case-control studies (www.ohri.ca/programs/clinical_epidemiology/oxford.htm). Items assessed included selection of cases and controls or cohorts, comparability, information concerning exposure/intervention, consideration for potential confounders, and percentage of withdrawals or dropouts. In particular, we paid close attention to the quality of the breastfeeding data, whether they were obtained prospectively or by retrospective recall, whether a distinction was made between exclusive and partial breastfeeding, and whether the duration of breastfeeding was reported. We also paid close attention to consideration of and appropriate adjustment for potential confounders.

We applied a three category summary grading system (A, B, C) to each study. This system defines a generic grading system that is applicable to each type of study design including randomized controlled trials, cohort, and case-control studies:

A (good)

Category A studies have the least bias and results are considered valid. A study that adheres mostly to the commonly held concepts of high quality including the following: clear description of the population, setting, interventions and comparison groups; clear description of the comparison groups; appropriate measurement of outcomes; appropriate statistical and analytic methods and reporting; no reporting errors; less than 20 percent dropout; clear reporting of dropouts; and appropriate consideration and adjustment for potential confounders.

B (fair/moderate)

Category B studies are susceptible to some bias, but not sufficient to invalidate the results. They do not meet all the criteria in category A because they have some deficiencies, but none of which are likely to cause major biases. The study may have suboptimal adjustment for potential confounders. The study may also be missing information, making it difficult to assess limitations and potential problems.

C (poor)

Category C studies have significant biases that may invalidate the results. The study either did not consider potential confounders or did not adjust for them appropriately. These studies have serious errors in design, analysis or reporting; have large amounts of missing information, or discrepancies in reporting.

For primary studies that had equivocal grading between moderate and poor, those studies were reviewed and graded again by different reviewers and consensus was reached after discussion among the reviewers. Lastly, it should be noted that the summary quality grading system evaluates and grades the studies within their own design strata. It does not attempt to assess the comparative validity of studies across different design strata. Thus, one should be cognizant of the study design when interpreting the methodological quality grade of a study.

Reporting of the Evidence

We reported each outcome separately in its own section in the results chapter. A brief explanation of the importance of the outcome is followed by a description of inclusion/exclusion criteria of studies examined that are specific to that outcome. A description of the relevant systematic review is followed by a description of the primary studies that were published after the latest search date of the systematic review. A summary table highlighted important findings. A summary conclusion regarding breast milk and that particular outcome is made in the last section. Conclusions were drawn only from studies of high or moderate (grade A or B) methodological quality. For dichotomous outcome, either the summary risk ratio or odds ratio is reported. For continuous outcome, the comparative difference in the actual measurement for that outcome is reported (e.g., IQ points, mm Hg of blood pressure). When only studies of C quality are available, we summarize the findings from those studies and explain the reasons for the “C” rating, but we do not draw conclusions from them.

Extracted data are compiled in evidence tables. The tables offer a detailed description of the studies that addressed each of the key questions. The tables (see Appendix C)^a provide detailed information about the study design, the sample size, the patient characteristics, the intervention and comparison group feeding methods, the followup, the major outcomes, and the methodological quality. In addition, for systematic reviews and meta-analyses, we reported the databases searched and for which time period, the number and the type of primary studies included, and the type of comparison addressed.

Summary tables succinctly report summary measures of the main outcomes evaluated. They include information regarding study design, intervention and comparison group, feeding methods, study duration or followup, sample size (subjects enrolled and analyzed in each arm), potential confounders, results of major outcomes, and methodological quality. These tables were developed by condensing information from the evidence tables. They are designed to facilitate comparisons and synthesis across studies. A methodological quality was assigned to each study as described previously.

^a Appendixes and Evidence Tables cited in this report are provided electronically at <http://www.ahrq.gov/clinic/tp/brfouttp.htm>.

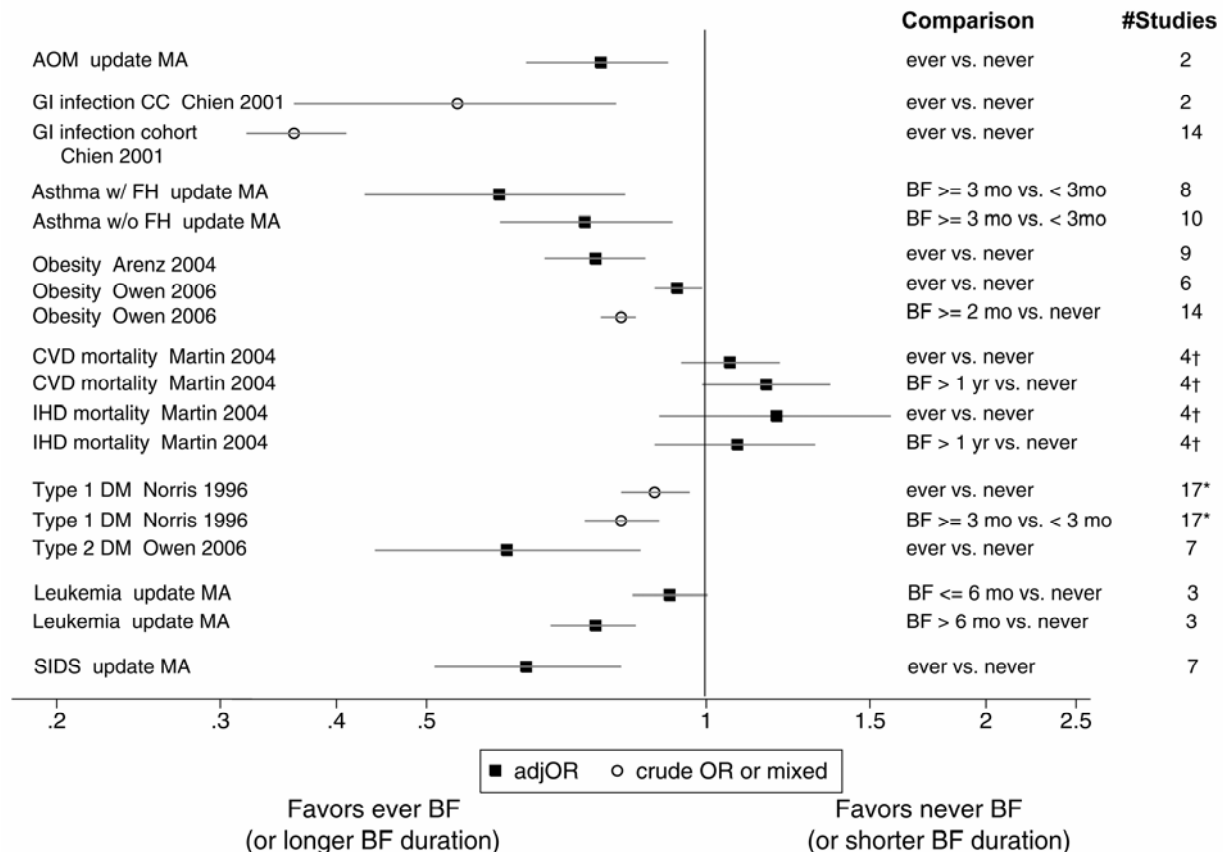
Chapter 3. Results

Overview

Twenty-three outcomes were analyzed in this report. We present three overall summary figures below to give the reader a quick overview of the results from the meta-analyses included in this report on the association of breastfeeding with health outcomes. We use the following rules to choose the results: (1) results from good or moderate quality meta-analyses, (2) the latest and/or highest quality meta-analysis is preferred if there are multiple meta-analyses addressing the same question, and (3) pooled adjusted estimate is preferred. Outcomes that did not have meta-analyses are not listed in these figures.

Three overall summary figures were created for term infant outcomes: Figure 1 for outcomes expressed as odds ratios or risk ratios comparing the different feeding groups; Figure 2 for the association between exclusive breastfeeding and infant outcomes; and Figure 3 for maternal outcomes expressed as odds ratios or risk ratios comparing the different feeding groups.

Figure 1. The relationship between breastfeeding and health outcomes in term infants - meta-analysis results

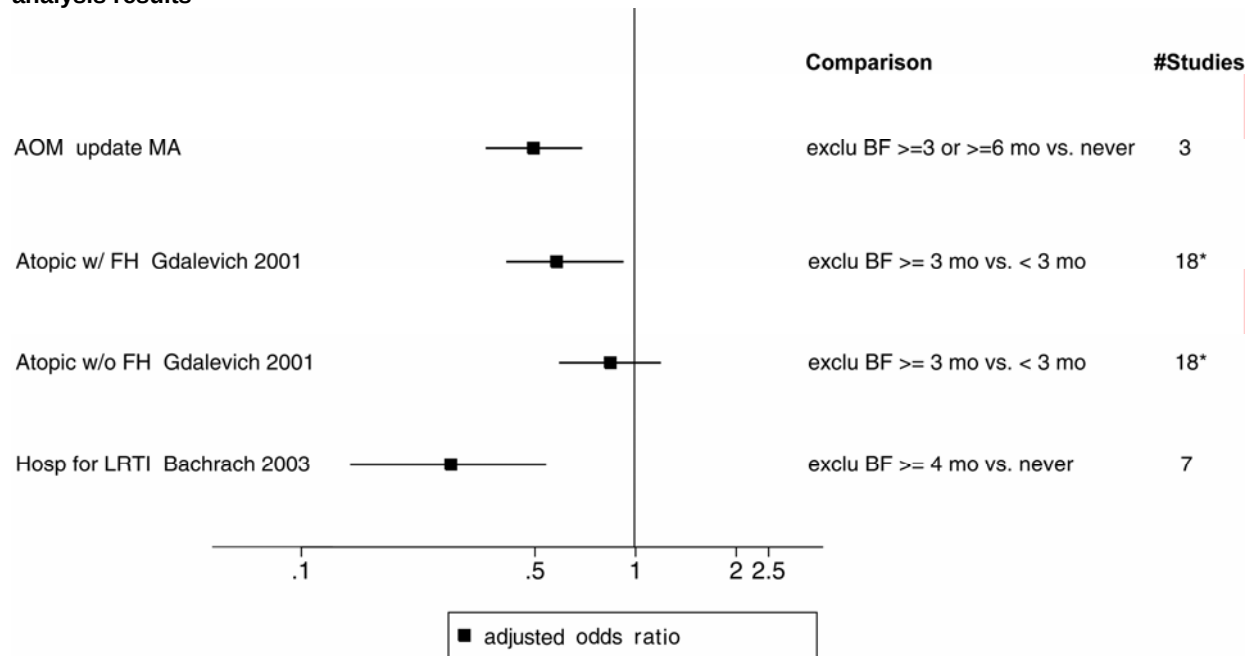


MA, meta-analysis; AOM, acute otitis media; GI, gastrointestinal; CC, case-control studies; FH, family history; CVD, cardiovascular disease; IHD, ischemic heart disease; DM, diabetes; adj, adjusted

*17 studies in total were included in Norris 1996 meta-analyses. The number of studies per comparison was not reported.

†Four historical cohort studies reported data on the relationship between breastfeeding and both CVD and IHD mortality.

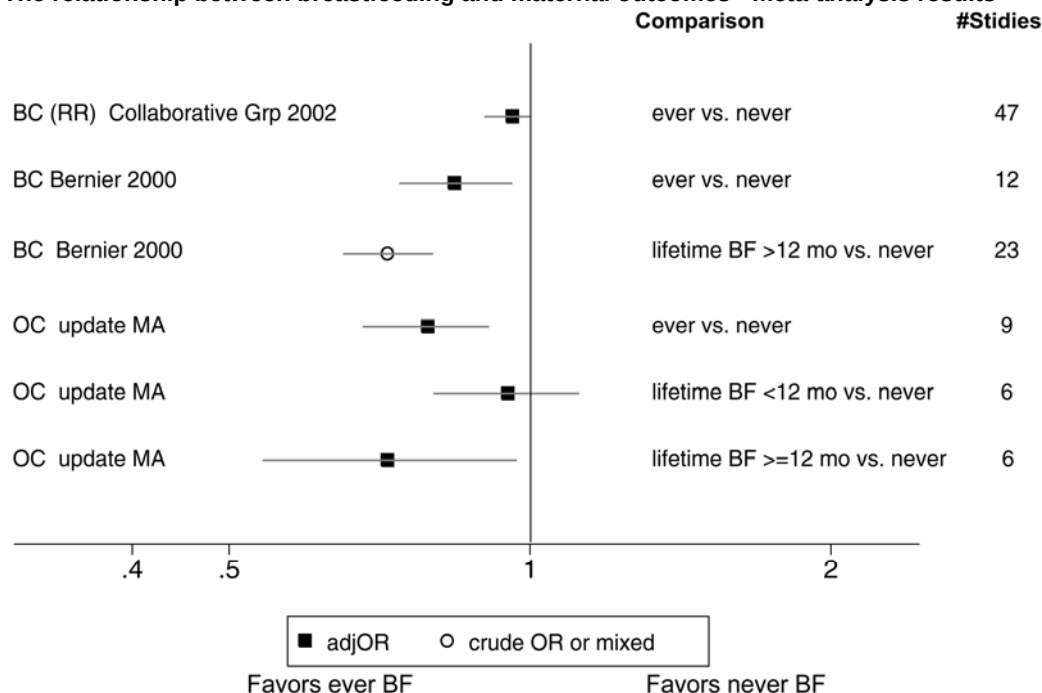
Figure 2. The relationship between exclusive breastfeeding and health outcomes in term infants - meta-analysis results



MA, meta-analysis; AOM, acute otitis media; FH, family history; Hosp, hospitalization; exclu, exclusive; LRTI, lower respiratory track infection

*18 studies in total were included in Gdalevich 2001 meta-analyses. The number of studies per comparison was not reported.

Figure 3. The relationship between breastfeeding and maternal outcomes - meta-analysis results



MA, meta-analysis; BC, breast cancer; RR, relative risk; OC, ovarian cancer; adj, adjusted

Literature Search Results

In the early phase of exploring the literature available for this report, a large number of published systematic reviews or meta-analyses were identified for various outcomes of interest. We screened over 300 abstracts for published systematic reviews and meta-analysis. We included all relevant systematic reviews and meta-analyses, although some of them included primary studies that were conducted in developing countries. A total of 29 systematic reviews or meta-analyses met our inclusion criteria. These systematic reviews or meta-analyses included a total minimum of 343 to a maximum of 494 unique primary studies (as some of the same studies were covered in multiple reviews). In this chapter, we summarized the findings and authors' conclusions from these systematic reviews and meta-analyses. If multiple systematic reviews and meta-analyses were available on the same outcome of interest, we also compared the differences and similarities between them.

We screened over 9,000 abstracts for potential relevant articles on the relationship between breastfeeding or breast milk feeding and various infant and maternal outcomes. After the initial screening, we categorized the abstracts according to the populations and outcomes of interest. For outcomes of interest that had previously been reviewed systematically, we excluded abstracts that were published before the search dates of previous systematic reviews or meta-analyses. For the remaining abstracts, we retrieved the corresponding full articles and applied additional inclusion or exclusion criteria tailored for the specific outcomes. For outcomes of interest that had not been reviewed systematically, we performed a new systematic review on all relevant primary studies that were conducted in developed countries.

Finally, a total of 43 unique primary studies on the relationship between breastfeeding and infant health outcomes, and a total of 43 unique primary studies on the relationship between breastfeeding and maternal health outcomes were included. Details of the inclusion and exclusion of abstracts and full article screenings were summarized in Figures 1 and 2. The full article inclusion or exclusion criteria were described in Chapter 2, and the additional inclusion or exclusion criteria tailored for the specific outcomes were described under each outcome section in this chapter.

Organization of Results

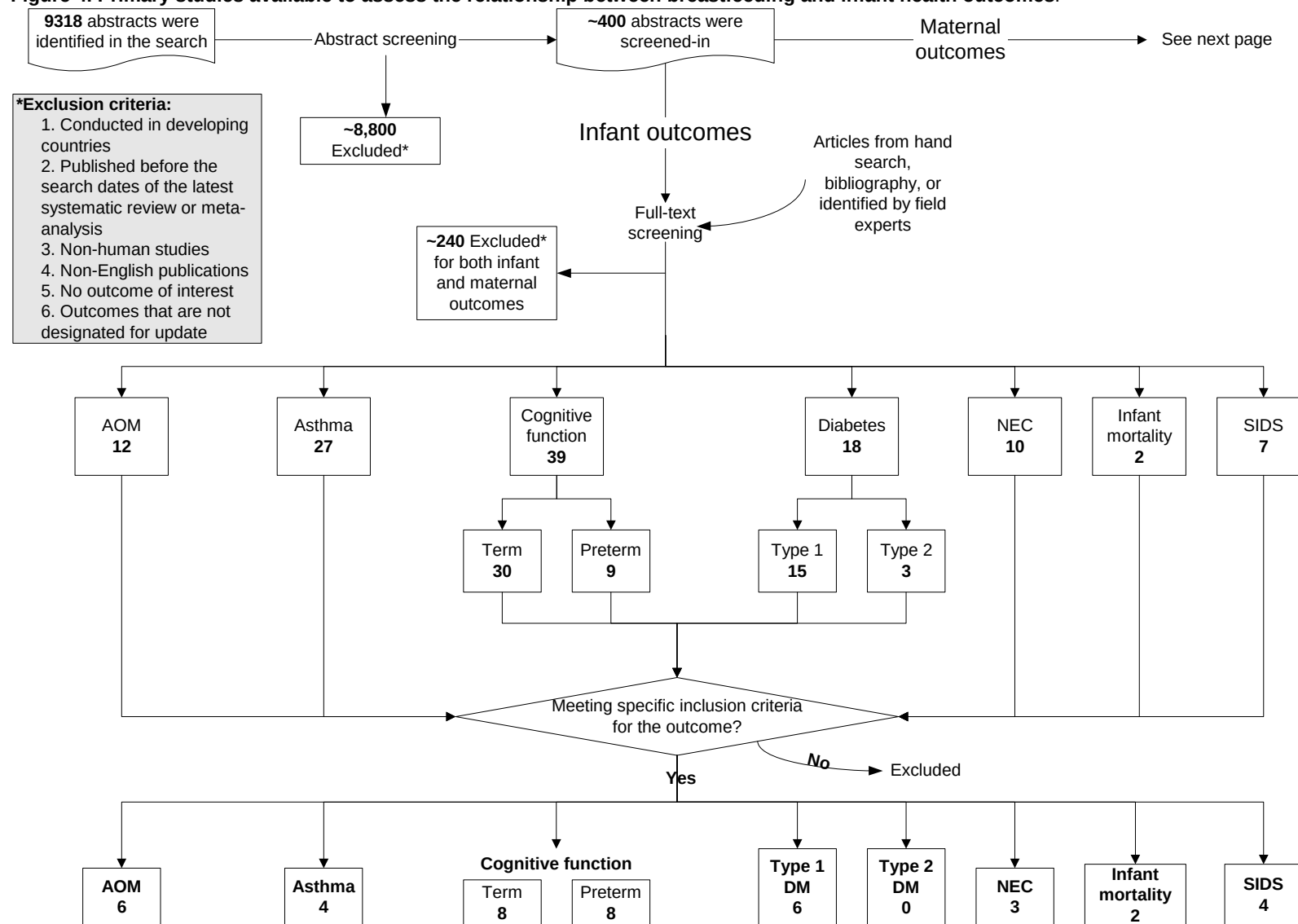
In this chapter, we grouped studies into one of the three parts of this report according to the target population – Part I: term infant outcomes, Part II: preterm infant outcomes, Part III: maternal outcomes.

In Part I, we summarized the results for health outcomes in term infants, including acute otitis media, atopic dermatitis, gastrointestinal infections, lower respiratory infection, asthma, cognitive development, obesity, cardiovascular diseases, cholesterols, blood pressure, cardiovascular mortality, type 1 and 2 diabetes, childhood leukemia, infant mortality, and sudden infant death syndrome.

In Part II, we summarized the results for necrotizing enterocolitis (NEC) and cognitive development in relationship to breast milk feeding in preterm infants, respectively.

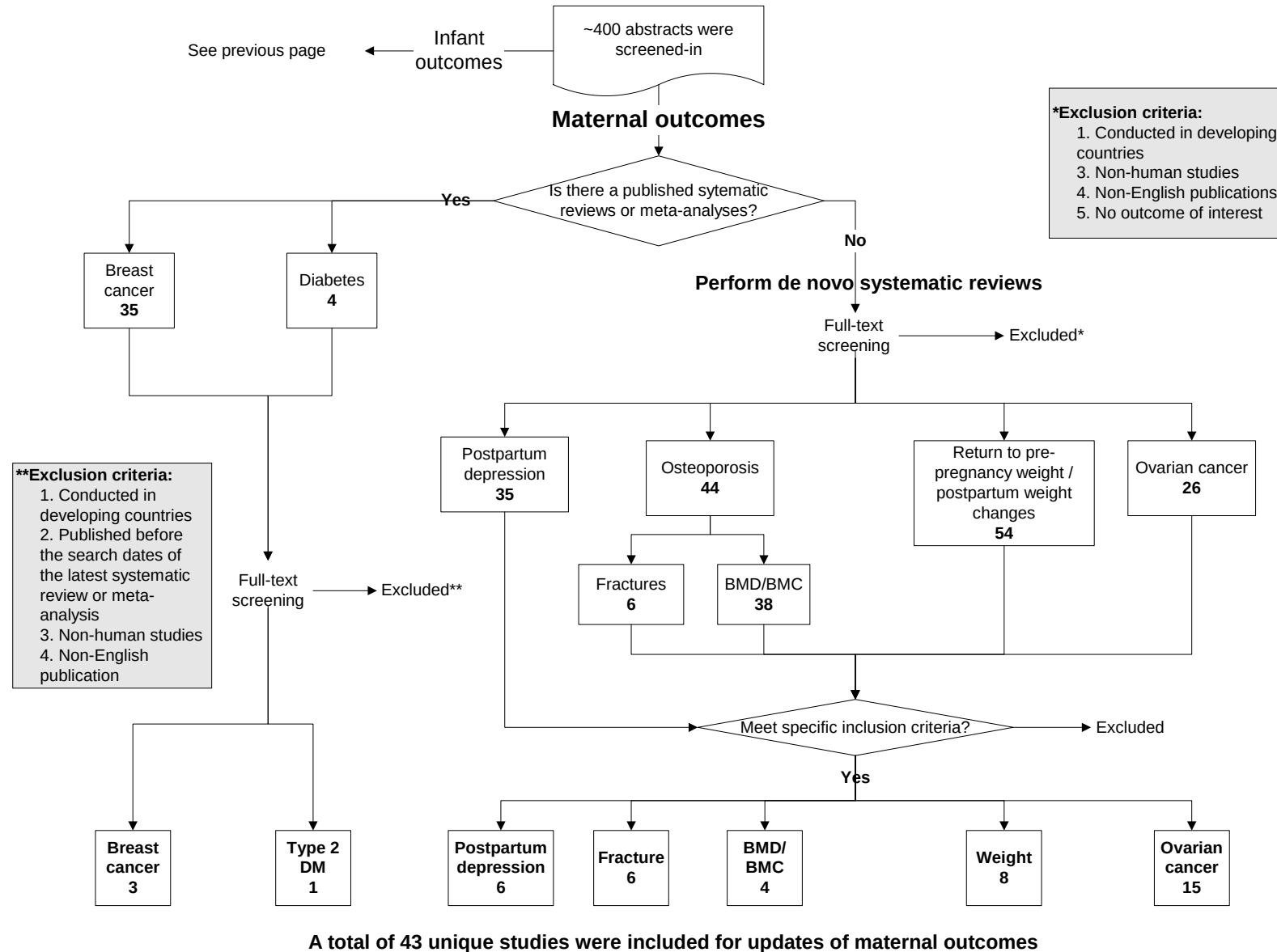
In Part III, we summarized the results for maternal health outcomes, including returning to pre-pregnancy weight, postpartum weight changes, maternal type 2 diabetes, osteoporosis, postpartum depression, breast cancer, and ovarian cancer.

Figure 4. Primary studies available to assess the relationship between breastfeeding and infant health outcomes.



A total of 43 unique studies were included for updates of infant outcomes

Figure 5. Primary studies available to assess the relationship between breastfeeding and maternal health outcomes.



Part I. Term Infant Outcomes

Relationship between Acute Otitis Media and Breastfeeding

Background

Acute otitis media (AOM) is a common childhood infection. It often begins with an upper respiratory tract infection. The viral infection predisposes the child to the development of AOM by causing eustachian tube dysfunction. The eustachian tube dysfunction enhances nasopharyngeal colonization with middle ear pathogens. The prevalence of a first attack of AOM in children under 1 year of age was estimated to be 44 percent.²⁵ Almost 70 percent of children under 6 years of age have had an episode of AOM.²⁵ Several risk factors have been identified for increasing the occurrence and recurrence of AOM.²⁶ There is a general consensus that breastfeeding protects against many infections, including AOM. Breast milk contains immunoglobulins with antibody activity against common bacteria such as *Haemophilus influenzae* and *Streptococcus pneumoniae*. It also contains components that interfere with the attachment of *Haemophilus influenzae* and *Streptococcus pneumoniae* to nasopharyngeal epithelial cells. The intermittent administration of milk with anti-adhesive substances into the nasopharynx of the nursing child may reduce the extent of colonization and protect against infection.²⁷

Commonly considered confounders in the studies of the relationship between breastfeeding and AOM were parental history of allergy, number of siblings, use of day care, maternal smoking, gender, ethnicity, and socioeconomic status.

We identified one meta-analysis of risk factors for AOM that aimed to further clarify possible means of preventing AOM in childhood.²⁸ Duration of breastfeeding was one of the factors examined. Since this meta-analysis, we identified six studies in seven publications that examined the relationship between breastfeeding and AOM.

Additional Methodological Comments

Uhari 1996 compared the risk of AOM among children who were breastfed for various durations – breastfeeding for 3 or more months versus less than 3 months, 6 or more months versus less than 6 months, and ever breastfeeding versus never breastfeeding.²⁸ The clinical and statistical heterogeneity of the studies included were not reported. The methodological quality of the meta-analysis was rated grade C, because there was no consideration for potential confounding and poor reporting of study characteristics (Table 1).

Since the meta-analysis was of poor quality, we decided to conduct a new meta-analysis combining adjusted odds ratios or risk ratios of AOM comparing breastfed infants with non-breastfed infants from the studies identified in Uhari 1996 and from the update search. We only included studies that reported the relationship between breastfeeding and the occurrence of AOM in infants without co-morbidities (e.g., cleft palate). Of the 10 studies included in Uhari 1996 meta-analysis, only three reported adjusted odds ratios or risk ratios of AOM comparing breastfed infants with non-breastfed infants.²⁹⁻³¹ Those studies are summarized in Table 2.

Studies Identified after the Systematic Review/Meta-analysis (Table 2)

A total of five cohort studies (in 6 publications)^{3,32-36} and one case-control study³⁷ that evaluated the relationship of breastfeeding and AOM published after the search dates of Uhari's meta-analysis met our inclusion criteria. We included only studies conducted in developed countries among children without co-morbidity. In cohort studies, subjects in the studies were followed from birth to a mean of 6 to 24 months. The number of subjects evaluated ranged from 289 to 15,113 at baseline. Four of the five cohort studies were of methodological quality grade B, while the other one was of grade C. In the case-control study, the children were between the ages of 3 and 7 years at the time of examination. A total of 179 AOM cases and 305 controls were analyzed. The methodological quality of the case-control study was rated C because the analysis did not control for potential confounding.

Studies varied in the definitions of breastfeeding and comparison groups. All studies were designed to evaluate a broad range of potential risk factors in AOM, except for Scariati 1997 in which the study specifically aimed to examine the relationship between breastfeeding and infections. We focused only on the relationship between breastfeeding and AOM. The studies also varied in their definitions of the disease conditions. In most studies, definitions of AOM were based on clinical features combined with otoscopic findings. The data were collected from medical records and confirmed by a physician in all studies except for Vernacchio 2004 and Scariati 1997 where a mother was asked to report ear infection in a list of diagnoses or symptoms.

All cohort studies adjusted for potential confounders and the case-control study did not. Confounding factors considered in the studies included gender, number of siblings, family day care, nursery day care, number of children in the home, maternal age, parental race or ethnicity, parity, maternal marital status, and parental smoking.

Breastfeeding was associated with a reduced risk of AOM compared to bottle-feeding after controlling for potential confounders across the five cohort studies, although some studies did not report a statistically significant result.

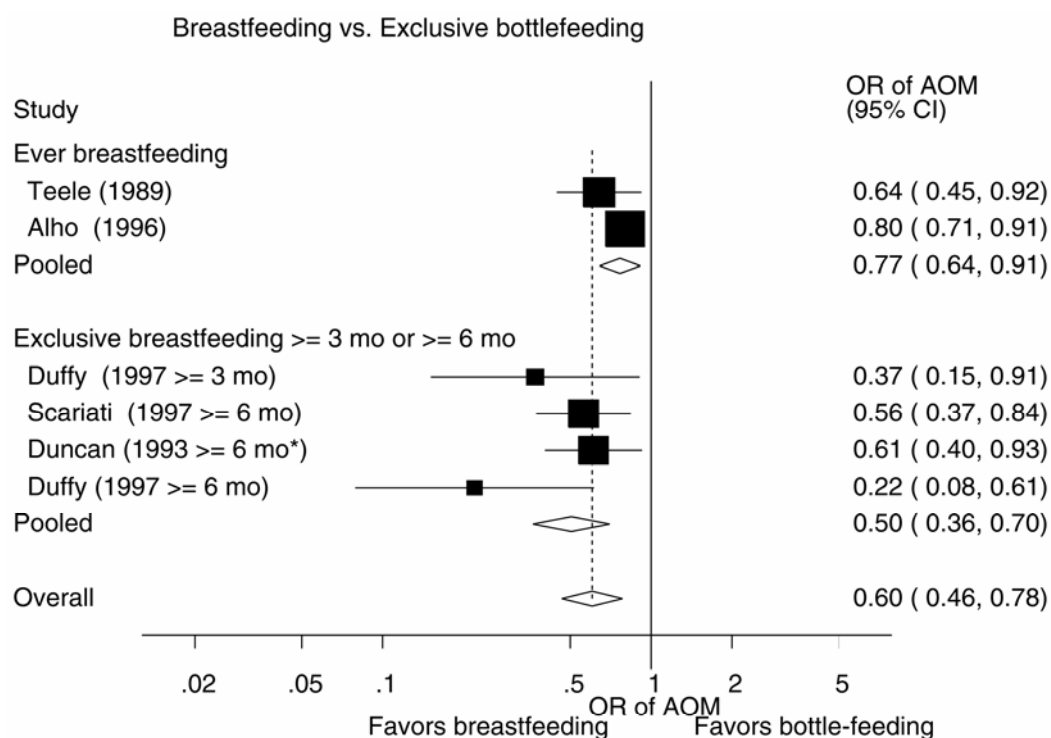
When exclusive breastfeeding for more than 3 or 6 months was compared with exclusive bottle- or formula- feeding, one study reported 28 percent³³ and another study reported 45 percent³ relative risk reduction in AOM. The case-control study did not show a significant difference in the risk of AOM between the children who were ever breastfed and those who did not. However, the analysis did not control for any potential confounding factors.

Meta-analysis. In addition to the five cohort studies in the update search, three additional cohort studies that reported adjusted odds ratios or risk ratios of AOM from Uhari 1996 were reviewed for possible inclusion into our meta-analyses. Studies were heterogeneous in the definitions of breastfeeding and comparison groups. In order to minimize the heterogeneity, we restricted our analyses to studies that reported an adjusted odds ratio or risk ratio of AOM comparing any definition of breastfeeding duration to exclusive bottle-feeding. Five studies that reported data from a total of six comparisons met the inclusion criteria for our meta-analyses.^{3,29,31,33,35} Using a random effects model, pooled data from five cohort studies of good and moderate methodological quality showed an adjusted odds ratio of AOM of 0.60 (95%CI 0.46-0.78) comparing breastfeeding to exclusive bottle-feeding. There is a significant difference in the risk reduction of AOM between the studies comparing exclusive breastfeeding with exclusive bottle-feeding and the studies comparing ever breastfeeding with exclusive bottle-feeding ($P<0.01$). Specifically, the pooled adjusted odds ratio of AOM was 0.77 (95%CI 0.64-

0.91), when comparing children who were ever breastfed with children who were exclusively bottle-fed in two studies. The pooled adjusted odds ratio was 0.50 (95%CI 0.36-0.70), when comparing children who were exclusively breastfed for at least 3 or 6 months with those who were exclusively bottle-fed for at least 3 or 6 months in three studies (providing 4 estimates).

The three studies that were excluded from our meta-analyses compared breastfeeding for more than 13 weeks, 4 months, and 6 months with breastfeeding for less than 13 weeks, 4 months, and less than 6 months, respectively.^{29,30,32} There was no significant association between the risk of AOM and breastfeeding for more than 13 weeks, while a significant risk reduction in AOM was found when comparing children who were breastfed for more than 4 months or 6 months with those who were breastfed for less than 4 months or 6 months, respectively.

Figure 6. Meta-analysis of the association between breastfeeding and the risk of AOM compared to exclusive bottle-feeding in cohort studies



* Exclusive breastfeeding ≥ 6 months vs. exclusive bottle-feeding and breastfeeding < 4 mo

Conclusion

The results from our meta-analyses of cohort studies of good and moderate methodological quality showed that breastfeeding was associated with a significant reduction in the risk of AOM. Comparing ever breastfeeding with exclusive bottle-feeding, the pooled adjusted odds ratio of AOM was 0.77 (95%CI 0.64 - 0.91). When comparing exclusive breastfeeding with exclusive bottle-feeding, either for more than 3 or 6 months duration, the pooled odds ratio was 0.50 (95%CI 0.36 - 0.70).

Table 1. Summary of systematic review/meta-analysis on the relationship between breastfeeding and acute otitis media (AOM)

Author Year	Study description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitations
Uhari 1996	MA of 2 case-control and 8 cohort studies in developed countries	Any measures of breastfeeding and comparators	ND	10 studies evaluated the risk of AOM and recurrence of AOM associated with breastfeeding. The diagnosis of AOM varied, but pneumatic otoscopy was used in the diagnosis. There was no restriction on study inclusion according to the diagnostic criteria used. 22 studies evaluating an array of risk factors for development and recurrence of AOM were included in the meta-analysis. Seven of the twenty-two studies specifically evaluated the association between breastfeeding and AOM, while another 3 studies included breastfeeding among other risk factors that were evaluated. Breastfeeding for at least 3 months reduced the risk of AOM (RR=0.87; 95%CI: 0.79-0.95) Children who were breastfed for 3 or more months had a non-statistically significant reduced risk of recurrent AOM compared with those who were breastfed less than 3 months (RR 0.69, 95%CI 0.46 - 1.03). When children who were breastfed for 6 months or more were compared to those who were breastfed for less than 6 months, a statistically significant reduced risk of recurrent AOM was found (RR=0.69; 95%CI: 0.49-0.97).	C No consideration for potential confounding, no restriction on inclusion according to the diagnostic criteria used for AOM; unclear how breastfeeding data was collected

AOM, Acute Otitis Media; SR, systematic review; MA, meta-analysis

Table 2. Summary table for cohort studies on the relationship between breastfeeding and acute otitis media

Author Year Country Design	Number of subjects (Baseline /Followup)	Definition of AOM	Mean duration Followup	Breastfeeding group (n)	Comparator group (n)	OR (95% CI)		Confounders / risk factors adjusted	Quality and limitations
						Crude	Adjusted		
Studies in update search									
Vernacchio 2004 USA Prospective cohort	15,113 /11,349	Mother selects ear infection from a list of diagnoses by physician in the past month	6 mo	BF at 6 mo (2,771)	No BF at 6 mo (8,558)	0.66 (0.59-0.74)	0.69 (0.61-0.78)	Race/ethnicity, gender, daycare attendance, cigarette smoking, income, access to care	B
Scariati 1997 US Prospective cohort	1,743 / 1410	Mother reported symptoms	7 mo	Exclusive BF at 6 mo (299)	Exclusive formula at 6 mo (891)	0.45	0.56, p<0.01	Gender, maternal education, occupation, smoking, households size and income, day care use	B
					1-57% mixed feeding (81)	0.56	0.63, p=0.03		
					58-88% mixed feeding (80)	0.67	0.71, p<0.05		
					89-99% mixed feeding (59)	0.83	0.83, p=NS		
Sassen 1994 Netherlands Prospective cohort	289 /232	Clinical diagnosis by physician. Tympanometry results not used.	23.6 mo	Overall effect before stopping BF (OR per month)	Up to 4 mo after stopping BF		0.92 (0.76-1.07)	Number of siblings, SES	B

Table 2. Continued

Author Year Country Design	Number of subjects (Baseline /Followup)	Definition of AOM	Mean duration Followup	Breastfeeding group (n)	Comparator group (n)	OR (95% CI)		Confounders / risk factors adjusted	Quality and limitations
						Crude	Adjusted		
Duffy 1997 Finland ^c Prospective cohort	Exclusive breast: 178 / 53 @ 3 mo / 28 @ 6 mo	Exclusive formula: 110 @ baseline	Exclusive formula: 110 @ baseline	Exclusive 3-mo BF	Exclusive 3- mo formula	0.72 (0.52-1.00)	0.72 (0.52-1.00)	Age at colonization, day care, maternal smoking	B
	Combined BF and formula: 18 @ baseline			Exclusive 6-mo BF	Exclusive 6- mo formula	0.55 0.34-0.87)	0.55 0.34-0.87)		
	Exclusive formula: 110 @ baseline			6-mo mixed feeding	Exclusive 6- mo formula	0.71 (0.53-0.96)	0.30 (0.19-0.48)		
	RR (95% CI)								
Alho 1996 (2 papers) Finland	2,512 / 825	>1 acute symptoms and one pneumatoscopic finding up to age 2 yrs	22 mo	Ever BF (nd)	Never BF (nd)	0.8 (0.7-0.9) ^e		Gender, number of siblings, atopy, age, season of birth, day care, parental smoking	C High dropout rate
Retrospective cohort				BF > 3 mo (735)	BF < 3 mo (90)	0.9 (0.8-1.0)			
Stenstrom, 1997 Sweden	AOM cases: 179 Control ^d : 305	> 5 episodes of AOM (Otosopic diagnosis) before age of 30 mo	N/A	Ever BF	Never BF	NS		None	C Unadjusted estimate only
Case-control study									
Studies in MA that reported adjusted OR / RR									
Duncan 1993 US	1220 / 1013	Diagnosed by experienced clinicians with an average of > 9 yrs in pediatric practice	12 mo	Exclusive BF ≥ 6 mo (154)	No BF and BF < 4 mo (465)	0.61 (0.40-0.92)		Parental history of allergy, number of siblings, use of day care, maternal smoking, gender, ethnicity, and maternal education	A
Prospective cohort				BF ≥ 4 mo, suppl 4-6 mo (199)		0.72 (0.54-0.95)			
				BF ≥ 4 mo, suppl <4 mo (200)		0.85 (0.74-0.97)			

Table 2. Continued

Author Year Country Design	Number of subjects (Baseline /Followup)	Definition of AOM	Mean duration Followup	Breastfeeding group (n)	Comparator group (n)	OR (95% CI)		Confounders / risk factors adjusted	Quality and limitations
						Crude	Adjusted		
Teele 1989 US Prospective cohort	1067 / 877	Effusion in one or both middle ears accompanied by ≥ 1 signs of acute illness ^b	12 mo	Ever BF (292)	Never BF (585)	≥ 1 episode of AOM: 0.83	≥ 1 episode of AOM: 0.64 (0.44-0.91) ≥ 3 episode of AOM: 0.51 (0.30-0.89)	Site of health care, season of birth, birth weight, gender, SES, number of siblings, sibling or parental history of infection, parental smoking	B
Howie 1990 Scotland Prospective cohort	750 / 617	Mother reported that infant experienced painful or discharging ear lasting for 48 hrs or more	12 mo	Exclusive BF ≥ 13 wk (89) or partial BF ≥ 13 wk (121)	Bottle feeders (246) and BF < 13 wk(161)	0.7 (0.5–1.0)	NS in the % infants experienced ear infection	Father's social class, maternal age, and parental smoking	B

BF, breastfeeding; excl, exclusive; ROM, recurrent otitis media; SES, socioeconomic status

^a Mean duration of exclusive BF (95% CI)

^b Signs of acute illness including earache, otorrhea, ear tugging, fever, irritability, lethargy, anorexia, vomiting or diarrhea

^c Proportion of infants in the individual categories of feeding method at different time cut-offs is not available

^d Matched to the case by age and gender

^e Values were estimated from figure. Statistically modeled to fit the whole sample.

Relationship between Atopic Dermatitis and Breastfeeding

Background

Atopic dermatitis is a common problem with an estimated lifetime prevalence in children of 10-20 percent.³⁸ Many studies have investigated the possible protective effect of breastfeeding on the development of atopic dermatitis. The results have been conflicting.^{16,38,39} Potential confounders considered in the studies included gender, socioeconomic status, family history of atopy, parental smoking, and presence of furry animals in the home.⁴⁰

Published Systematic Review/Meta-Analysis (Table 3)

We identified one systematic review/meta-analysis that examined the relationship between breastfeeding and the development of atopic dermatitis.⁴⁰ The methodological quality of this systematic review/meta-analysis was rated grade A.

Using the MEDLINE database from 1966 to 2000, Gdalevich 2001 identified 18 prospective studies from developed countries that qualified for inclusion in the review. Sample size of the studies ranged from 17 to 991. A total of 4,158 participants were included. Mean followup duration was 4.5 years. Study selection criteria included term infant, restriction of maternal recall of the child's feeding history limited to the first 12 months of the infants' life, the breastfeeding group in the study were exclusively breastfed for at least 3 months, blinding of the feeding history during outcome assessment, strict diagnostic criteria of atopic dermatitis provided by the authors, and control for confounding variables like socioeconomic status and family history of atopy.

Using a fixed effect model, the overall summary odds ratio for the development of atopic dermatitis was 0.68 (95% CI 0.52 - 0.88) in those subjects with at least 3 months of exclusive breastfeeding versus subjects without 3 months of exclusive breastfeeding. When the analysis was restricted to those studies with positive family history of atopy, the odds ratio was 0.58 (95%CI 0.41 - 0.92). When the analysis was restricted to those studies without a family history of atopy, the odds ratio was 0.84 (95%CI 0.59 - 1.19).

The systematic review did not make a distinction between atopic dermatitis of infancy (under 2 years of age) and persistent or new atopic dermatitis at older ages. However, in a primary analysis, the data were stratified according to different durations of followup (because the diagnosis of atopic dermatitis in patients younger than 2 years of age are sometimes attributed to symptoms of infectious origin and breastfeeding may have a protective effect against infections). The summary odds ratio in the group with less than 2 years of followup was 0.74 (95%CI 0.61 – 0.90), whereas the summary odds ratio in the group with 2 or more years of followup was 0.78 (95%CI 0.62 – 0.99).

The authors of the review concluded that there was a substantial protective effect of breastfeeding against atopic dermatitis in children with a family history of atopy.

Conclusion

Available evidence from one well-performed systematic review/meta-analysis on full term infants in developed countries suggests that exclusive breastfeeding for at least 3 months was

associated with a reduction in the risk of atopic dermatitis in those subjects with a family history of atopy.

Table 3. Summary of systematic review/meta-analysis on the relationship between breastfeeding and the risk of developing atopic dermatitis

Author Year	Studies description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA
Gdalevich 2001 Term infants 4,158 (17-991)	18 prospective cohort studies in developed countries	≥ 3 mo of exclusive breastfeeding vs. without ≥ 3 mo of exclusive breastfeeding	SES, gender, family history of atopy, parental smoking, and presence of furry animals at home; also considered disease in < 2 yr of age vs. ≥ 2 yr (to sort out the possible confounding protective effect of breastfeeding for skin rash that is infectious in etiology which might be confused with the diagnosis of atopic dermatitis)	Overall OR 0.68 (95% CI 0.52 - 0.88, fixed effect model); When restricted to those with positive family history, OR 0.58 (95%CI 0.41 - 0.92) When restricted to those without family history, OR 0.84 (95%CI 0.59 - 1.19) Authors' conclusion: "There is a substantial protective effect of breastfeeding against atopic dermatitis in children with a family history of atopy."	A

SR, systematic review; MA, meta-analysis; SES, socioeconomic status

Relationship between Gastrointestinal Infections and Breastfeeding

Background

Gastrointestinal infections are common in infants and children. Rate of diarrheal disease in US was estimated to be 1.1 episodes per person-year in children less than 5 years old.⁴¹ Many studies have investigated the possible protective effect of breastfeeding on the development of gastrointestinal infections. A previous review of diarrhea morbidity in both developed and developing countries reported that the risk of diarrhea in infants who did not receive breast milk were 3.5 to 4.9 times higher than infants who had exclusive breastfeeding in the first 6 months of life.⁴² Factors like secretory IgA, oligosaccharides, lactoferrin and others available in breast milk may protect the infant from various infections through passive immunity.¹ In vitro and in vivo binding studies have demonstrated that fucosylated glycans in breast milk inhibit binding by campylobacter jejuni, stable toxin of enterotoxigenic Escherichia coli, and major strains of calciviruses (e.g., noroviruses (also known as Norwalk-like viruses)) to their target host cell receptors.⁴³ One report suggests that glycoprotein lactadherin found in breast milk protects against rotavirus infection.⁴⁴ Socioeconomic status and child care variables (e.g., home versus day care, degree of crowding at home) are thought to be important confounding factors in observational studies on gastrointestinal infection and breastfeeding. As the aforementioned review also included studies before 1950, it would be instructive to review the more recent literature and assess the relationship between breastfeeding and gastrointestinal infections in infants from developed countries.

Published Systematic Review/Meta-Analysis (Table 4)

We identified one systematic review/meta-analysis that examined the relationship between breastfeeding and the development of gastrointestinal infections in children less than 1 year of age from developed countries.⁴⁵ The methodological quality of this systematic review/meta-analysis was rated grade B.

Using the MEDLINE database from 1966 to 1998, and supplementing the results with searches of the bibliographies of the primary and review articles, Chien 2001 identified 16 studies from developed countries that qualified for inclusion in the review. There were 12 prospective cohort studies totaling 5,473 subjects, 2 retrospective cohort studies totaling 504 subjects, and 2 case-control studies with 331 pairs of subjects. For the review, gastrointestinal infection was defined to be “any illness associated with vomiting, change in consistency or frequency of stools, or isolation of a known enteropathogenic bacterial or viral agent.” For the final data analysis, infant feeding practices were dichotomized into two groups: exclusive breastfeeding and partial/mixed feeding, or exclusive artificial feeding.

Results were conflicting. Nine of 16 studies (56 percent) yielded a statistically significant protective effect of breastfeeding on gastrointestinal infections. Majority of the studies suffered from methodological deficiencies. Four studies fulfilled criteria of controlling for detection bias, analyses of confounders, having a clear definition of infant feeding practices and infectious outcomes. Three of the studies reported breastfeeding was protective against non-specific gastrointestinal infection. The fourth study reported that differences in feeding practice did not

affect the attack rates of rotavirus gastroenteritis. The potential confounders examined in these studies included infant sex, race, maternal education, family living standards, marital status, paternal social class, and/or parental smoking. Even though these studies adjusted for potential confounders, the actual quantitative adjusted odds ratio or risk ratio were not reported.

The authors of the systematic review stated that “it was not possible to pool the adjusted relative measures of association” in the cohort studies reviewed. Using a fixed effect model, the summary crude odds ratio of the 14 cohort studies for the development of gastrointestinal infection was 0.36 (95% CI 0.32, 0.41; heterogeneity $P < 0.01$); the summary odds ratio of the two case-control studies was 0.54 (95% CI 0.36 - 0.80; heterogeneity, $P = 0.35$).

Conclusion

Available evidence from three primary studies that controlled for potential confounders suggests that breastfeeding is associated with a reduction in the risk of non-specific gastrointestinal infection during the first year of life in infants from developed countries. However, a summary adjusted effect estimate taking into account potential confounders could not be determined because not all the studies that adjusted for potential confounders provided usable quantitative data for meta-analysis.

Addendum

During the final phase for the preparation of this report, we were alerted to a recent primary study on diarrheal disease in infants in 1990s England that provided relevant quantitative data. This case-control study of 304 infants (167 cases and 137 controls) showed that the infants who were breastfeeding had a reduced risk of diarrhea compared to infants who were not breastfeeding (OR 0.36, 95% CI 0.18 to 0.74, $P = 0.005$). The result was adjusted for age, sex, social class, contact with person in and outside household, and other factors. This study also reported that the protective effect of breastfeeding did not persist beyond 2 months after cessation of breastfeeding.⁴⁶

Table 4. Summary of systematic review/meta-analysis on the relationship between breastfeeding and gastrointestinal (GI) infection

Author Year Population	Studies description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA
Chien 2001 Prospective cohorts: 5,473 Retrospective cohorts: 504 Case-control: 331 pairs	12 prospective cohort studies, 2 retrospective cohort studies; 2 case-control studies	Either exclusive breastfeeding and partial/mixed feeding or exclusive artificial feeding	Maternal education, single or dual parenting, family living standards, infant sex, race, maternal parity, marital status, paternal social class, maternal age, parental smoking	Conflicting results on the effect of breastfeeding on GI infection: 9/16 studies (56%) yielded a statistically significant protective effect of breastfeeding on GI infections; Majority of studies suffered from methodological deficiencies; 4 studies fulfilled criteria of controlling for detection bias, analyses of confounders, having a clear definition of infant feeding practices and infectious outcomes; 3 of these studies reported breastfeeding was protective against non-specific GI infection; Unadjusted pooled estimate of the cohort studies: OR 0.36 (95% CI 0.32, 0.41; heterogeneity P<0.01) Unadjusted pooled estimate of 2 case- control studies: OR 0.54 (95% CI 0.36, 0.80; heterogeneity P=0.35)	B

SR, systematic review; MA, meta-analysis

Relationship of Hospitalization Secondary to Lower Respiratory Tract Diseases in Infancy and Breastfeeding

Background

Respiratory infection is the most common medical problem among infants and children. Each year in the United States, three percent of all infants are hospitalized with moderate to severe respiratory infection.⁴⁷ Severe lower respiratory tract diseases may increase the risk of childhood asthma. Viral infections, especially respiratory syncytial virus (RSV) infection, are the most common cause of lower respiratory tract disease in developed countries. RSV infection occurs most frequently between 2 and 8 months of age. The risk factors associated with rates of respiratory illness include age, smoke exposure, day care, race/ethnicity, family size, education, and socioeconomic status. We identified one meta-analysis published in 2003 that compared the risk of hospitalization for respiratory diseases in healthy full term infants who were breastfed with those who were not.⁴⁸

Published Systematic Review/Meta-Analysis (Table 5)

Bachrach 2003's meta-analysis included seven cohort (five prospective and two retrospective) studies that evaluated the relationship between breastfeeding and hospitalization risk secondary to respiratory diseases.⁴⁸ Primary studies published before April 2002 and evaluated healthy full-term infants less than 1 year of age from developed countries were included. Only data from exclusive breastfeeding of at least 2 months or any breastfeeding totaling 9 months or more were included. Exclusive breastfeeding was defined as little or no formula feeding. Studies on cystic fibrosis and allergic conditions were excluded. Unpublished data were also examined. The primary outcome variable was hospitalization for lower respiratory tract disease secondary to bronchiolitis, asthma, pneumonia, empyema, and infections due to specific agents (e.g., RSV). The meta-analysis was restricted to first hospitalization. The analysis evaluated the risk of hospitalization in 3,201 breastfed subjects and 1,324 non-breastfed subjects.

The meta-analysis used a random effects model. There was an overall 72 percent reduction in the risk of hospitalization in infants who were exclusively breastfed for 4 or more months compared with those who were formula-fed (summary relative risk 0.28, 95% CI 0.14 - 0.54). There was no change in summary relative risk in the subgroup analyses of studies that also reported relative risk adjusted for the effects of smoking or socioeconomic status. There was no statistical heterogeneity across the studies. The authors performed sensitivity analyses to assess the appropriateness of combining studies and found no changes in summary risk. In addition, the number needed to treat (NNT) was calculated to be 26. This implies that at least 26 infants will have to breastfeed exclusively for 4 or more months to prevent one infant from hospitalization secondary to respiratory diseases. The methodological quality of the meta-analysis was rated grade A.

Conclusion

The meta-analysis showed an overall 72 percent reduction in the risk of hospitalization secondary to respiratory diseases in infants who were exclusively breastfed for 4 or more months

compared with those who were formula-fed. This finding remained statistically significant after adjustment for potential confounders. There could be differences across studies with regard to duration of followup, diagnosis of respiratory disease, and other factors. However, there was no statistically significant heterogeneity detected across the studies. Taking into account the methodological quality of the meta-analysis and the consistent findings reported, we conclude that breastfeeding for 4 or more months is associated with a reduction in the risk of hospitalization secondary to lower respiratory tract diseases.

Table 5. Summary of systematic review/meta-analysis on the relationship between breastfeeding and lower respiratory tract disease

Author year Population	Study description	Intervention /Comparator	Confounders considered	Outcomes	Quality of SR/MA
Bachrach 2003 Infants <1 yr of age Exposed: 3,201 Unexposed: 1,324	MA of 7 observational studies	Breastfeeding exclusively ≥ 2 mo or ≥ 9 mo total of any (including mixed feed)	Smoking or socioeconomic status	Summary RR of 0.28, 95% CI 0.14- 0.54 There was an overall 72% reduction of hospitalization among infants who were exclusively breastfed for ≥ 4 mo compared with those who were formula-fed.	A

SR, systematic review; MA, meta-analysis

Relationship between Asthma and Breastfeeding

Background

Asthma is a bronchial disorder accompanied by breathing difficulties, wheezing, coughing, and production of thick mucus. Triggering factors have been attributed to foreign substances, tobacco smoke or pollutants, exercise, infection, and emotional stress. An increase in the prevalence of asthma has been reported in some countries during the second half of the 20th century.⁴⁹ In 2002-2003, about six percent of the children in the United States had an asthma attack (www.cdc.gov/nchs/data/has/has05.pdf). In addition, the hygiene hypothesis (a lack of early childhood exposure to infectious agents increases susceptibility to allergic diseases^{50,51}) has been proffered as a possible explanation for some of the increase in incidence and prevalence of asthma, but that hypothesis remains a matter of debate.⁴⁹ Potential confounders considered in the studies of breastfeeding and asthma include age, socioeconomic status, family history of atopy, and parental smoking.

Findings from primary studies on the relationship between breastfeeding and the development of asthma have been conflicting. One meta-analysis published in 2001 aimed to examine this relationship.⁵² We identified three relevant prospective longitudinal cohort studies and one followup study (to one of the primary studies in the meta-analysis) published subsequent to the meta-analysis.⁵³⁻⁵⁶ Three of the four studies met the inclusion criteria of the original meta-analysis and an updated meta-analysis stratified by family history of asthma was conducted.

Published Systematic Review/Meta-Analysis (Table 6)

Gdalevich 2001 conducted a meta-analysis that included 12 prospective observational studies with 8,183 term infants followed for a mean of 4.1 years. Study inclusion criteria included subjects from developed countries, exclusive of breastfeeding for at least 3 months, blinding of diagnosing physician to feeding status, and maternal recall of child's feeding history of not more than 12 months. Exclusive breastfeeding was defined as having no substitutes or additions. Critical appraisal of the studies was conducted based on the standards suggested by Kramer.¹⁶ Potential confounders including age, socioeconomic status, family history of atopy, and parental smoking were controlled for by means of multivariate analysis. The outcome of asthma was diagnosis by a physician.

Meta-analysis of the 12 studies with followup of 2 or more years reported a summary odds ratio of 0.70 (95% CI 0.60 - 0.81), suggesting an association of breastfeeding and a reduction in the risk of the development of asthma. Subgroup analysis reported that children with a family history of asthma or atopy benefited more from breastfeeding (OR 0.52, 95% CI 0.35 - 0.79) in the risk reduction of the development of asthma, compared with children who did not have a family history of asthma or atopy. There was no statistical heterogeneity in the studies. Sensitivity analysis indicated the exclusion of any one study did not change the overall results. Methodological quality for the meta-analysis was rated grade A.

Studies Identified after the Published Systematic Review/Meta-Analysis (Table 7)

We identified three prospective cohort studies, including one that had both prospective and retrospective analyses, and one followup publication examining the relationship between breastfeeding and risk of the development of asthma.⁵³⁻⁵⁵ Methodological quality of these four studies ranged from grade C to grade A.

Subjects selected in these studies ranged from healthy full-term infants to any live births from mainly population-based samples. Sample sizes at followup ranged from 1,037 to 4,964 with zero to 31 percent dropouts or withdrawals. Verification of asthma varied from parental confirmation through questionnaires to clinic assessment and physician diagnosis. Although all four studies adjusted for confounding, the only variable common to all studies was maternal smoking, either during pregnancy or after birth.

Two of the studies were population-based, one was hospital-based, and one study was done in a large health maintenance organization (HMO). Wright 2001 and Kull 2004 looked at exclusive breastfeeding whereas Burgess 2006 did not define breastfeeding but acknowledged that exclusive breastfeeding was in practice difficult to verify. Another issue is that there were no gold standards for the diagnosis of asthma in population studies. To confirm the diagnosis of asthma, Kull 2004 relied on clinical examinations as well as blood sampling and pulmonary function tests. Burgess 2006 used data from a questionnaire that asked about medications, symptoms, hospitalization, and family history to formulate the diagnosis. Sears 2002 based the diagnosis on pulmonary function tests with the addition of maternal description of symptoms. Wright 2001 and Kull 2004 appeared to have the most rigorous criteria for diagnosis of asthma by considering recent symptoms, results from testings, and confirmation by a physician. Two studies gathered feeding data post-delivery by recall, first study at 6 months,⁵³ and the second study at 2 months with another followup at 12 months.⁵⁴ Two studies relied on objective data for feeding history; Sears 2002 relied on data from nursing program records,⁵⁵ and Wright 2001 relied on breastfeeding data from clinic visits.⁵⁶ The methodological quality of the studies by Kull, Wright, and Sears were rated grade B, and Burgess was rated grade C.

Study findings. Kull 2004 followed 3,384 newborns for the first 4 years of life and reported that there was a statistically significant association of exclusive breastfeeding for 4 months or more and a reduction in the risk of the development of asthma (OR 0.72, 95%CI 0.53 - 0.97).⁵⁴ Furthermore, subgroup analysis showed that the association was stronger for children whose parents did not have a history of allergic diseases.

Sears 2002 enrolled 1,037 subjects at age 3 years and a retrospective record review found that approximately half were breastfed and half were not (533 breastfed, 504 not breastfed).⁵⁵ The subjects were followed prospectively until 21 or 26 years of age. The study found an increased risk of asthma in those subjects who were breastfed compared with those who were not for all time points assessed. Length of breastfeeding duration had no protective effect against asthma. Family history of asthma did not significantly affect these results. Even though “exclusive breastfeeding” rate was reported in the study, the authors acknowledged that it was common practice for the hospital staff to feed the newborns with formula for the first few nights post-delivery to allow the mothers to sleep.

Wright 2001 is a followup publication to one of the primary studies from the Gdalevich 2001 meta-analysis. This study reported that there was no association between duration of exclusive breastfeeding and asthma for 926 children by age 13 years followed since birth, except in those

children who were atopic and whose mothers had a history of asthma.⁵⁶ This was the only study to report that a family history of asthma is an effect modifier of breastfeeding in the increased risk of the development of asthma in children.

Burgess 2006 investigated 4,964 children at 14 years of age concerning their history of asthma (“yes” or “no”) as reported by their mothers. This study found that there was no significant relationship between the duration of breastfeeding and the prevalence of asthma. The rates of asthma were the same for any given rates of breastfeeding regardless of the maternal history concerning asthma.⁵³ Data on feeding were collected at 6 months after birth.

Updating the Previous Meta-Analysis

We performed an update meta-analysis stratified by family history of asthma using the random effects model including three of the four recent publications.^{53,54,56} The fourth study did not qualify for inclusion because it did not have a comparison group of at least 3 months duration of breastfeeding. In the subgroup analysis of those children with a positive family history of asthma, two studies (Wright⁵⁶ and Kull⁵⁴) met the inclusion criteria set by Gdalevich. One of the studies reported a very large adjusted odds ratio (OR 8.7, 95%CI 3.4 – 22.2) in a followup of 13 year olds, compared with the other studies.⁵⁶ Sensitivity analyses were conducted to examine the source of heterogeneity by including or excluding this study. Excluding Wright, a history of exclusive breastfeeding for more than 3 months was associated with a reduction in the risk of asthma (OR_{adj} 0.60; 95%CI 0.43 - 0.82), compared with no breastfeeding. Including Wright, there was no longer a statistically significant association between a history of breastfeeding and the risk of asthma (OR_{adj} 0.81, 95%CI 0.41 – 1.60). This result suggests that the heterogeneity can be explained by a single study. Compared with the other studies in the analyses, the age of followup was 13 years in Wright 2001, while it ranged from 2 to 9 years in the other studies.

In the analysis of those children without a family history of asthma, the addition of the two subsequent studies (Kull⁵⁴ and Burgess⁵³) did not alter the statistically significant association of breastfeeding and the reduction in the risk of asthma (OR 0.73, 95%CI 0.59 – 0.92) compared with the original results reported by Gdalevich. It should be noted that Burgess reported only the unadjusted odds ratio, stating that the adjusted was “minimally altered”.

Figure 7. Meta-Analysis of prospective cohort studies of the association between asthma risk and breastfeeding ≥ 3 months for children with positive family history of asthma or atopy (excluding Wright 2001)

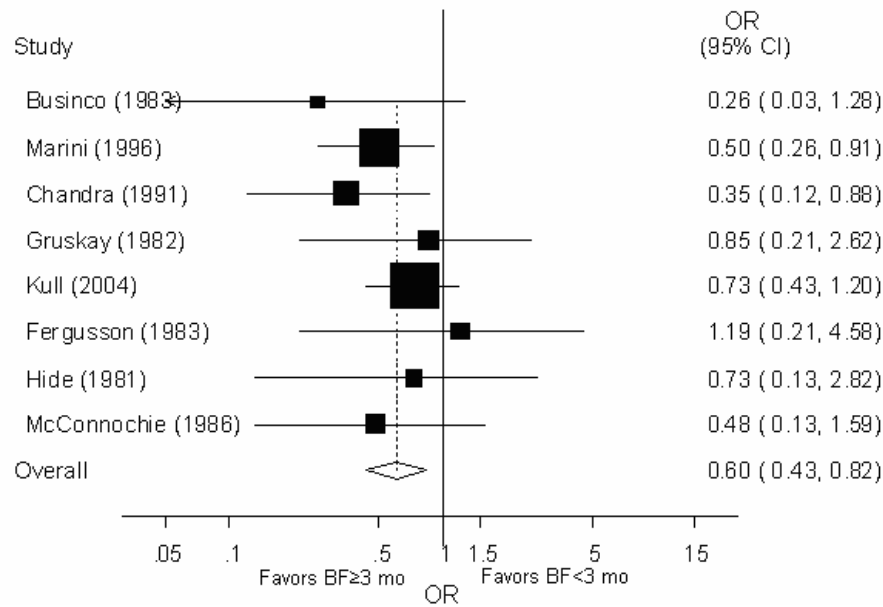


Figure 8. Meta-Analysis of prospective cohort studies of the association between asthma risk and breastfeeding ≥ 3 months for children with positive family history of asthma or atopy (including Wright 2001)

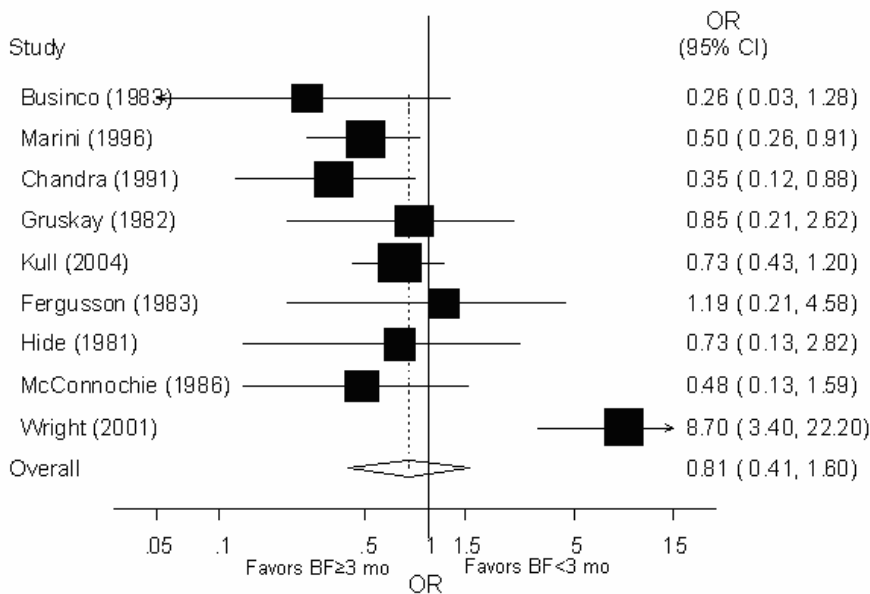
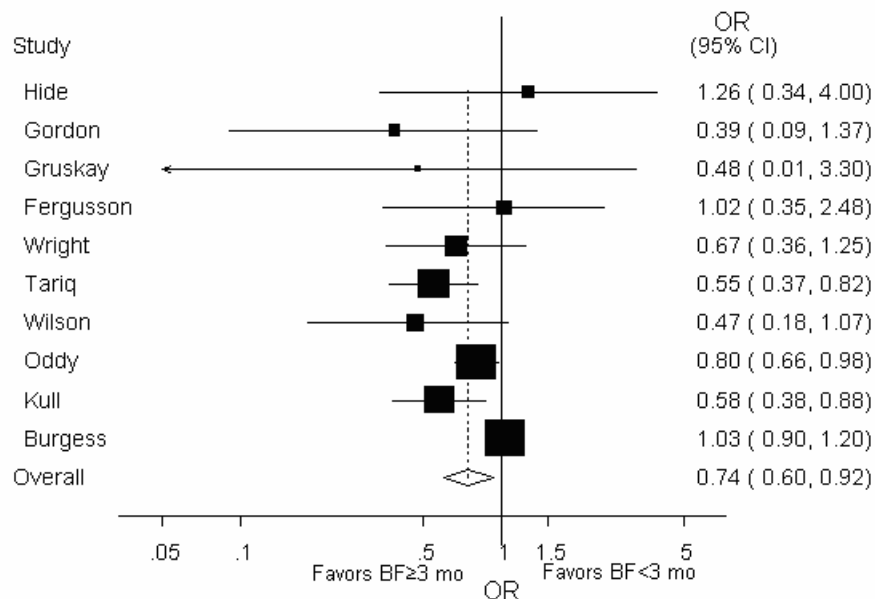


Figure 9. Meta-Analysis of prospective cohort studies of the association between asthma risk and breastfeeding ≥ 3 months for children without family history of asthma or atopy



Conclusion

A well-performed meta-analysis from 2001 concluded that breastfeeding was associated with a reduction in the risk of developing asthma. This association was stronger in those subjects with a positive family history. However, three new primary studies and one followup study reported conflicting results. With our new meta-analyses including three of these studies, it is clear that there remained an association between breastfeeding and a reduction in the risk of asthma in those subjects without a family history of asthma. This association was also found in subjects under 10 years of age with a positive family history of asthma. It is unclear whether this association changes for older children. It should also be noted that the fourth study, which did not qualify for inclusion in our new meta-analyses, reported an increase in asthma risk with increased duration of breastfeeding in those subjects with a maternal history of asthma. Further studies concerning the effect of a family history of asthma on long-term outcome of asthma is warranted.

Table 6. Summary of systematic review/meta-analysis on the relationship between breastfeeding and asthma

Author Year Population	Studies description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitations
Gdalevich 2001 Term infants	MA of 12 prospective studies in developed countries	≥ 3 mo exclusive breastfeeding vs. without ≥ 3 mo of exclusive breastfeeding	SES, gender, family history of atopy, parental smoking, and presence of furry animals at home	≥ 2 years followup, breast feeding had protective effect against asthma, summary OR 0.72 (95% CI 0.62 – 0.84) Subgroup analysis of children from family with asthma or atopy showed higher protective effect of breast feeding OR 0.52 (95% CI, 0.35 - 0.79) Subgroup analysis of children from family without asthma or atopy, OR 0.73 (95% CI, 0.62 - 0.86)	A

SR, systematic review; MA, meta-analysis

Table 7. Summary of prospective cohort studies on the relationship between breastfeeding and asthma

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment of asthma	Confounders adjusted for	Outcomes	Quality and limitations
Kull 2004 Sweden Enrolled: 4,089 Follow-up: 3,384 All newborns from catchments area of Stockholm	Prospective, longitudinal cohort with questionnaires at child's age of 2 months, 1, 2, and 4 years. Clinic assessment at 4 years.	Exclusive breastfeeding was dichotomized with the 25 th percentile as the cutoff point (<4 months and ≥4 months) and as a 3- level categorical variable (0-2, 3-4, and ≥5 months)	At age 4 years, 4 episodes of wheezing during the last 12 months or at least 1 episode of wheezing during the same period if the child was receiving inhaled steroids	Maternal age, maternal smoking during pregnancy or at 2 months of age, and heredity of allergic diseases	- Overall exclusive breastfeeding for ≥ 4 month associated with lower risk than those who breastfed <4 months, adjOR 0.72 (0.53-0.97) - Children with no heredity of allergic diseases was associated with lower asthma risk, adjOR 0.58 (0.38-0.88) - Children with heredity of allergic diseases had adjOR 0.73 (0.43-1.20)	B No blinding
Sears 2002 New Zealand Enrolled: 1,037 Followup:1,037 Population- based cohort from single hospital center	Both prospective and retrospective cohort study design enrolled at age 3 and data collected retrospective and prospective, followed prospectively up to 21 years	Breastfeeding categorized by duration: Not breastfed Breastfed > 4 weeks	Asthma as reported by child or parent	SES, birth order, sheepskin use in infancy, maternal smoking	Breastfeeding at least 4 weeks compared with no to less than 4 weeks associated with higher adjusted OR of asthma, 2.40 (1.36-4.26)	B
Wright 2001 USA Enrolled: 1,246 Followup: 1,043 Healthy newborns from large HMO	Prospective, longitudinal newborn cohort with questionnaires at 2, 3, 6, 9, 11, 13 years; Additional data collection by MD health surveillance	Exclusive breastfeeding categorized by duration: never breastfed breastfed <4 months breastfed >4 months	Physician diagnosed asthma, wheezing or asthma symptoms reported ≥ 2 questionnaires from ages 6 to 13 years	Maternal education, smoking status in 1 st year, sex, ethnicity, 2 or more siblings at home or day care use versus neither first 6 months, paternal asthma	- Nonsignificant relationship between asthma and exclusive breastfeeding duration at 13 years - Children with maternal history of asthma: exclusive breastfeeding associated with asthma, adjusted OR 8.7 (3.4-22.2)	B No explanation for withdrawals/dropouts; discrepancies in numbers of exclusive breastfeeding; no blinding

Table 7. Continued

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment of asthma	Confounders adjusted for	Outcomes	Quality and limitations
Burgess 2006 Australia Enrolled: 7,223 Followup: 4,964 Singleton term deliveries from public clinic of University Hospital	Prospective, longitudinal cohort with questionnaires at birth, 6 months, 5 and 14 years.	Breastfeeding categorized by duration: Never breastfed < 3 weeks 3 – 6 weeks 7 weeks – 3 months ≥ 4 months	From questionnaire: episodes of asthma in past 6 months, frequency of asthma meds, asthma- related sick days from school, asthma-related hospital admissions	Maternal asthma, paternal asthma, smoking early and late pregnancy, frequency of coughs and cold first 6 months, annual family income	Breastfeeding at least 4 months not associated with asthma compared to no breastfeeding, unadjusted OR 1.03 (0.9-1.2)	C Adjusted OR not reported, dropout/withdrawals > 30%, significant differences between completers and noncompleters, no definition or description of breastfeeding, discrepancy for % males in table.

SES, socioeconomic status

Relationship between Cognitive Development and Breastfeeding in Term Infants

Background

Many studies have examined the relationship between breastfeeding and cognitive development. Results have been conflicting. Many of them did not have a clear definition of breastfeeding or breast milk exposure. Different cognitive assessment tools were used. Outcomes were measured anywhere from less than 2 years of age to adulthoods. Confounders commonly considered in these studies were socioeconomic status, maternal education, birthweight, gestational age, birth order and gender.⁵⁷ Jain et al. considered SES and quality and quantity of stimulation of the child (including social interactions) to be crucial confounders.⁵⁸ These authors did not consider maternal or paternal intelligence, marital status, number of children, and maternal age to be crucial because to “some degree, they are markers of socioeconomic status, and are not clearly related to both feeding method and intelligence independent of socioeconomic status”. Three systematic reviews from 1999 to 2002 have tried to either establish methodological standards to assess the observational studies or adjust for covariates in pooled analysis. Since the last systematic review by Jain et al. in 2002,⁵⁸ there have been eight prospective cohort studies on healthy term infants that examined the relationship of breastfeeding to some aspects of cognitive development.

Additional Methodological Comments

We searched for systematic reviews on breastfeeding and cognitive development using terms like “cognitive”, “neurodevelopment”, “intelligence”...etc. Using similar terms, we searched MEDLINE and CINAHL in January and April of 2006, respectively, for additional primary studies published after 2000. We also identified additional articles based on reviews of the bibliographies cited in the relevant retrieved studies from the search and from suggestions by the reviewers for this report. For primary studies, qualifying study designs included prospective cohort and case-control studies; only studies from developed countries were included. For the healthy term infants, subgroups like infants of diabetic mothers and small-for-gestational age infants were not included. For preterm infants, no subgroups were excluded. Studies concerned exclusively with individual breast milk factor supplements (e.g., long chain polyunsaturated fatty acids, nucleotides) were not included. There was no restriction on timing of and the tools used for cognitive assessment. Only data on cognitive outcomes were extracted, data from motor, psychomotor, or visual development were not extracted.

Published Meta-Analyses/Systematic Reviews (Table 8)

Jain 2002 identified 40 relevant publications (30 birth cohorts, five RCTs, five school registry cohorts, and three case-control studies) from 1929 to 2001. Most of them studied term infants. A few studies included only preterm infants in their analyses. Each study was assessed according to a set of eight clinical epidemiological standards: study design, target population, sample size, quality of feeding data (suitable definition and duration of breastfeeding, and appropriate timing and source of feeding data), whether studies controlled for socioeconomic

status and stimulation of the child, whether observers of the outcome were blind to feeding status, whether a standardized individual test of general intelligence at older than 2 years of age was used, and whether studies reported an effect size. Two studies on term infants met all the proposed methodological standards. One concluded that the effect of breastfeeding on intellect was significant (4.6 points in IQ at age 3 years in children who breastfed compared with those who bottle-fed),⁵⁹ and the other concluded that the effect on cognitive performance was statistically non-significant.⁶⁰

Drane 2000 examined 24 studies from 1966 to 1998.⁵⁷ Twenty-one included term infants, two included low birthweight infants, and one involved small-for-gestational age infants. Each study was assessed according to a set of three methodological standards: definition of outcome, correct classification of type of infant feeding, and control of potential confounding variables (socioeconomic status, maternal education, birthweight, gestational age, birth order, gender). Five of the 24 studies met all three methodological standards. These studies indicated an advantage in IQ to breastfed infants in the range of two to five points for term infants and eight points for low birth weight infants.

Interestingly, of the two studies that were identified by Jain 2002 to have met all the proposed methodological reporting standards, the feeding data from one did not meet the feeding data reporting standard put forth in the review by Drane 2000. The study at issue was Wigg 1998, in which the following statement could be found: "Feeding methods (i.e., breast, bottle-fed or mixed) and duration of breastfeeding in infancy were recorded at age 6 months by the trained research nurse".⁶⁰ Jain 2002 interpreted that study to have met their requirement of adequate reporting of feeding data ("whether infants received breast milk exclusively or with supplemental formula or other foods"), while Drane 2000 interpreted that study not to have adequately distinguished between partial and exclusive breastfeeding. The other study that met all the methodological standards according to Jain 2002 was published in 1996,⁵⁹ but it was not included in Drane 2000's review; reason of which was not readily apparent.

Anderson 1999 examined 11 studies from 1978 to 1995 that controlled for at least five covariates (breastfeeding duration, sex, maternal smoking history, maternal age, maternal intelligence, maternal education, maternal training, paternal education, race or ethnicity, socioeconomic status, family size, birth order, birth weight, gestational age, and childhood experiences) and presented unadjusted and adjusted results.⁶¹ Eight studies included term infants; three studies included only preterm infants. The results were combined in a meta-analysis. The adjusted (fixed effects) pooled mean difference was 3.16 points (standardized effect estimate of cognitive developmental mean score, 95% CI 2.35 - 3.98) in favor of the breastfeeding group. Low birthweight infants showed larger differences (5.18 points in cognitive developmental score; 95%CI 3.59 - 6.77) than did normal birth weight infants (2.66 points; 95%CI 2.15 - 3.17). Other methodological quality of the studies was not assessed.

Studies Identified after the Published Meta-Analysis/Systematic Review in Term Infants (Table 9)

Prospective cohort. One secondary analysis of a prospectively collected data set⁶² and seven prospective cohort studies ranging from 44 to 3880 term subjects reported on the relationship between breast milk feeding and some aspects of cognitive development since 2002.⁶³⁻⁷¹

None of the cohort studies made a clear distinction between partial and exclusive breastfeeding. Two studies stated to have a prospective design, but the breastfeeding information

was collected retrospectively. Subjects were breastfed from less than 3 weeks to 12 months. Three studies reported the proportion of subjects who breastfed for more than 6 to 7 months, they ranged from 9 percent to 62 percent.

Bayley Mental Development Index Scale (MDI) was the cognitive assessment tool for subjects under 2 years of age. Peabody Picture Vocabulary Test (PPVT-R) and Wechsler Preschool and Primary Scale of Intelligence (WPPSI-R) were used in subjects under 7 years of age. Wechsler Adult Intelligence Scale (WAIS), Raven's standard progressive matrices, standard reading, verbal, and mathematical reasoning tests were used in adults up to 27 years of age. Time of assessment ranged from 1 to 27 years.

Der 2006 analyzed the database from the US National Longitudinal Survey of Youth 1979 (NLSY79).⁶² This database has information on the participants and their offsprings. The study reported that the adjusted effect of breastfeeding on Peabody individual achievement test (standardized mean of 100 and SD of 15) at 14 years was reduced to +0.52 from +4.7 after adjustment for maternal IQ, education, age, family poverty, home stimulation, and birth order. Further analysis of 332 pairs of siblings discordant for breastfeeding status found non-significant difference between groups in both status and duration of breastfeeding. Meta-regression of nine unique studies (including the data from NLSY79) reported an advantage of breastfeeding of 0.16 cognitive points after controlling for maternal IQ and other confounders. This study was rated good methodological quality (grade A).

Der 2006 also combined its estimate with the estimate from the only other sibling analysis in the literature to date (Evenhouse 2005,⁷² this study did not qualify for inclusion in this review because it was an analysis of data obtained from a cross-sectional design study). Evenhouse 2005 analyzed the database from the National Longitudinal Study of Adolescent Health (Add Health) 1994 (This database oversamples low-income, African-American, and Hispanic children and provided information on an abbreviated Peabody Picture Vocabulary Test results. There were 523 pairs of siblings with different breastfeeding history in this data set). The combined estimate from NLSY79 and Add Health was 0.025 ($P=0.54$) for breastfeeding status and 0.04 ($P=0.271$) for duration of breastfeeding.

Of the six prospective cohort studies of moderate (grade B) methodological quality, five reported an advantage in cognitive development in subjects who breastfed. Specifically, Lawlor 2006 reported that the adjusted score (for sex, parental characteristics, birthweight, and perinatal characteristics) in Raven's standard progressive matrices at age 14 years showed a mean difference of 6.79 (95%CI, 5.33-8.26) in less than 4 months of breastfeeding versus never breastfeeding (compared to an unadjusted mean of 8.20).⁷¹ Quinn 2001 reported that the mean PPVT-R at 5 years for those breastfed for at least 6 months was 8.2 points (95%CI, 6.5-9.9) higher for females and 5.8 points (95%CI, 4.1-7.5) higher for males when compared to those never breastfed. PPVT-R was adjusted for birthweight, poverty, maternal education, maternal age, time in daycare or preschool, the number of children in the household at 5 years, English speaking background in parents, and infant stimulation.⁷⁰ Similarly, Oddy 2003 and 2004, reported that the PPVT-R score at 6 years was 3.56 point higher for children breastfed more than 6 months compared with children never breastfed ($F=8.59$, $P=0.003$).^{68,69} The result was adjusted for gender, gestational age, maternal age and education, parental smoking, and the presence of older siblings. Mortensen 2002 reported that the duration of breastfeeding was associated with significantly higher scores on the verbal, performance, and full scale WAIS at 27 years.⁶⁷ With regression adjustment for parental social status and education, single mother status, mother's height, age, and weight gain during pregnancy, cigarette smoking during 3rd trimester, number of

pregnancies, estimated gestational age, birth weight, birth length, and indices of pregnancy and delivery complications, the mean full scale WAIS were 99.4, 101.7, 102.3, 106.0, and 104.0 for breastfeeding durations of < 1 month, 2 to 3 months, 4 to 6 months, 7 to 9 months, and > 9 months, respectively ($P=0.003$). GomezSanchiz 2004 reported that the Bayley MDI at 24 months was 4.3 points higher in those breastfed more than 4 months compared with those breastfed less than 4 months after multiple linear regression adjusting for parental IQ.⁶⁶ Angelsen 2001 reported that adjustment for differences in maternal intelligence reduced the odds ratio of having a low IQ score among children who were breastfed for <3 months compared to ≥ 6 months from 2.8 (95%CI 1.4-5.3) to 1.5 (95%CI 1.0-2.1).⁶⁴

Conclusion

One well-performed sibling analysis and prospective studies that controlled specifically for maternal intelligence demonstrated that there is either little or no evidence for an association between breastfeeding and cognitive performance in children. It is clear that maternal intelligence is a major confounder in the studies on relationship between breastfeeding and cognitive development. For those studies that still reported a significant effect after adjustment for maternal intelligence, residual confounding from other factors like different home environments cannot be ruled out. Many studies controlled for socioeconomic status and maternal education but not specifically for maternal intelligence. It is clear that maternal intelligence should be controlled for separately from socioeconomic status and maternal education in any studies of breastfeeding and cognitive development. As cautioned by Der et al., “The generalizability of the results presented here must be considered carefully. This study and the others included in the meta-analysis are all based on samples from developed countries. Generalization of the findings beyond these and similar societies would be unwise. We have also excluded premature and low birthweight infants for whom the effect may be different.”⁶²

Table 8. Summary of systematic reviews/meta-analyses on the relationship of breastfeeding and cognitive development

Author Year Population	Studies description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitations
Jain 2002 50 to >11,000 Term and preterm	30 birth cohorts; 2 RCTs; 5 school registry cohorts; 3 case-control studies	Breastfeeding, breast milk, or choice to breastfeed	SES (parental education, occupation, income or any combination), quality and quantity of stimulation of the child (including social interactions)	Quality of feeding data evaluated by 4 criteria (exclusive breastfeeding or not, timing of feeding data collection, source of feeding data, duration of breastfeeding): 9 articles met all 4 criteria (8 full-term cohorts). 22 studies used an appropriate measure of cognition. Only 2 studies met all the methodological standards. One study concluded that "any beneficial effect of breastfeeding on cognitive development is quite small in magnitude", another study found that children who were breastfed had mean IQ scores 4.6 points higher than those never breastfed after controlling for socioeconomic status and other factors. Among the studies that controlled for socioeconomic status and stimulation/interaction of the child, three concluded that breastfeeding promotes cognitive development, and four did not.	A
Drane 2000 N in individual studies not provided in the review Term and preterm	24 studies (1 RCT and 23 cohorts) met inclusion criteria: birth 1960-98, English language, examined cognitive development	Actual breastfeeding intervention in individual studies not described in the review	SES, maternal education, birthweight, gestational age, birth order, gender	5 studies met the 3 methodological standards (operational definition of cognitive outcome and outcome measure using standardized tests; correct classification of type of feeding: measure breastfeeding as a continuous variable or at least as a 3-level categorical variable (exclusive breast or formula-fed; partial breastfed) and control for potential confounding variables. The only RCT did not find statistical significant difference in Bayley MDI in infants fed solely on a diet of donor breast milk compared with infants fed solely on a diet of term formula. Advantages in IQ as measured by the WISC-R, Bayley and McCarthy were observed in 4 cohort studies. In term infants, effects on IQ in the range of 2-5 points (0.2-0.3 SD) were found.	B

Table 8. Continued

Author Year Population	Studies description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitations
Anderson, 1999 7,081 Term and low birthweight	11 cohort studies with results adjusted for covariates	Breastfeeding vs. formula feeding (no detailed information)	breastfeeding duration, sex, maternal smoking, age, intelligence, education, training, paternal education, race or ethnicity, SES, family size, birth order, birth weight, gestational age, and childhood experiences	Adjusted (fixed effects) pooled mean difference of standardized effect estimate of cognitive developmental score was 3.16 (95% CI 2.35, 3.98) in favor of breastfeeding; An average adjusted benefit from breastfeeding of 5.18 points in cognitive developmental score was obtained for low-birth weight children across 6 available observations.	B No additional quality assessments of primary studies

MDI, Bayley Mental Development Index; SR, systematic review; MA, meta-analysis; SES, socioeconomic status

Table 9. Summary of studies on the relationship of breast milk feeding and cognitive development

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment tools for cognitive development	Confounders adjusted for	Outcomes	Quality and limitations
<i>Database analysis of a prospective cohort study</i>						
Der 2006 US Database from the US national longitudinal survey of youth 1979 (NLSY79) and children of the women in the survey, Full term infants	Database analysis of a prospective study; sibling pairs analysis, and meta-analysis; Cognitive test administered to children between 5 yr and 14 yr biennially from 1986 to 2002; For meta-analysis, only included studies that quantified the effect of breastfeeding on cognitive ability after controlling for parental intelligence	Breastfeeding history obtained within a year of birth in most cases	Peabody individual achievement test (PIAT) was administered to children	Adjustment for variables associated with breastfeeding in the survey, home environment (HOME-SF), child demographics, maternal characteristics	PIAT (all outcomes standardized to a mean of 100 and SD of 15) Unadjusted effect of breastfeeding +4.7 compared to non-breastfeeding (3,161 mothers, 5,475 children, 16,744 assessments); after adjustment for maternal Armed Forces Qualification Test (AFQT) score, education, age, family poverty, HOME stimulation score, and birth order, the difference became +0.52 (P=0.149) 332 pairs of sibling discordant for breastfeeding status and 545 discordant for duration of breastfeeding, difference between groups (status) = -0.63 (P=0.506); (duration) = -0.13 (P=0.866) Meta-regression of 9 unique studies (including the data from NLSY79): an advantage of breastfeeding of 0.16 after controlling for IQ and 8 additional confounders. Combined data from NLSY79 and sibling analysis study by Evenhouse 2005 (see separate extraction): estimate 0.025 (P=0.54) for breastfeeding status and 0.04 (P=0.271) for duration of breastfeeding	A No details regarding breastfeeding history

Table 9. Continued

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment tools for cognitive development	Confounders adjusted for	Outcomes	Quality and limitations
<i>Prospective cohort</i>						
Lawlor 2006 Australia Enrolled 7,223 Followup 3,794 (>98% term)	Enrolled subjects at birth and did Peabody Picture Vocabulary test at 5 yr and Raven's standard progressive matrices (non- verbal reasoning or general intelligence) at 14 yr, most of the data reported were from age 14 (see Quinn 2001 for 5-yr data)	Classified by never, < 4 mo, ≥ 4 mo (obtained from mothers at the 6-mo followup)	Raven's standard progressive matrices at 14 years	Sex, parental characteristics (maternal age, ethnicity, education, paternal education, family income, gravity, maternal smoking), labor, Apgar scores, birthweight, height, BMI	At age 14 yr, Never breastfed = 694, <4 mo = 1372, ≥ 4 mo = 1,606 All parental characteristics were related to offspring IQ score. Unadjusted scores at age 14 showed a mean difference of 4.43 (95%CI 3.09 to 5.77) in <4 mos breastfeeding vs never breastfeeding; 8.20 (95%CI 6.89 to 9.49) in ≥ 4 mo breastfeeding vs. never breastfeeding (P<0.001) Adjusted scores at age 14 yr (N=3,099) showed a mean difference of 4.07 (95%CI 2.61 to 5.53) in <4 mo breastfeeding vs. never breastfeeding; 6.79 (95%CI 5.33 to 8.26) in ≥ 4 mo breastfeeding vs. never breastfeeding (P<0.001) Family income, parental education and breastfeeding explained 7.5% of the variation in intelligence at age 14 yr. Loss to followup was selective, those subjects were more likely to have mothers who were from poorer social backgrounds, lower education, and younger; regression analysis repeated using Heckman's sample selection bias adjustment with maternal age, parental education, and family income as the selection variables; results of these regression models did not differ from those who had followup.	B Large drop out

Table 9. Continued

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment tools for cognitive development	Confounders adjusted for	Outcomes	Quality and limitations
Quinn 2001 Australia Enrolled 7,357 Followup 3,880 Singleton; children with major neurological abnormalities and those for whom the data were incomplete were excluded	Enrolled women attending antenatal clinic; data collected at baseline, after birth, 6 mo and 5 yr	No distinction was made between partial or complete breastfeeding; classified by never, < 3 wk, 3 to < 7 wk, 7 wk to < 4 mo, 4 mo to < 6 mo, or ≥ 6 mo	Peabody Picture Vocabulary Test (PPVT- R) was administered at 5 years.	Birth weight, poverty, maternal education, maternal age, time in daycare or preschool, number of children in the household at 5 years	Breastfeeding ≥ 6 mo 103.6 (SD 13.1) No breastfeeding 94.2 (SD 14.1) There was a significant trend towards increasing PPVT-R with increased duration of breastfeeding (P=0.0000) Before adjustment, the mean for those breastfed ≥6 mos was 10.9 points (95%CI 9.3, 12.5) higher for females and 7.5 points (95% CI 5.9, 9.1) higher for males when compared to those never breastfed. After adjustment, the mean for those breastfed ≥6 mos was 8.2 points (95%CI 6.5, 9.9) higher for females and 5.8 points (95%CI 4.1, 7.5) higher for males when compared to those never breastfed.	B Large drop out

Table 9. Continued

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment tools for cognitive development	Confounders adjusted for	Outcomes	Quality and limitations
Oddy, 2003, 2004 (same data) Australia 2,393 at birth, 1,444 at 6 yr, 1,371 at 8 yr 37-39 wk gestation	Comparing no breastfeeding, <4 mo, 4-6 mo, and >6 mo of breastfeeding	90% of children were breastfed at some point, 28% were breastfed for >6 mo, solid foods were introduced at ≤6 mo in 88% of infants.	Peabody Picture Vocabulary Test (PPVT- R) was administered at 6 y/o and a Performance subtest (Perceptual organization WISC- Block Design) at 8 y/o	Gender, gestational age, maternal age and education, parental smoking, and the presence of older siblings.	Both verbal IQ and performance scores increase with increasing maternal education combined with a longer duration of breastfeeding, with the most profound effect of breastfeeding occurring in the highest education groups ($P<0.005$). In the lower education groups, these trends were less consistent. Before adjustments for covariates, children breastfed for >6 mo had mean verbal IQ scores 6.44 points higher and Block Design scores that were 1.13 points higher than children never breastfed. After adjustment for covariates, there was an association between duration of breastfeeding and verbal IQ with a 3.56 point advantage for children breastfed >6 mo compared with children never breastfed ($P=0.003$). The adjusted association of full breastfeeding with the Performance subtest was not significant ($P=0.223$). There was an association of low PPVT-R and non-English maternal language; children of non-English speaking parents ($n=100$) were excluded from further analyses.	B Large drop out

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment tools for cognitive development	Confounders adjusted for	Outcomes	Quality and limitations
Mortensen 2002 Denmark Enrolled 9,125 (4668 males) 973 had WAIS; 2,280 had BPP & BF data	Prospective cohort; but breastfeeding information collected retrospectively at 1- year exam	Divided into 5 BF groups: ≤ 1 mo 32%; 2-3 mo 32%; 4-6 mo 24%; 7-9 mo 9%; > 9 mo 3%	2 sub-cohorts; one (50% male) took the WAIS at 27 yr; the other (100% male) took the draft board intelligence test: Borge Priens Prove (BPP) at 18.7 yr	Parental social status and education, single mother status, mother's height, age, and weight gain during pregnancy, cigarette during 3 rd trimester, number of pregnancies, estimated gestational age, birthweight, birth length, and indices of pregnancy and delivery complications	Breastfeeding duration (mo) <1 2-3 4-6 7-9 > 9 Unadjusted Adjusted WAIS Unadjusted Adjusted BPP 98.1 101.3 103.3 108.2 102.8 99 102 102 106 104 P=0.003 for overall F 36.9 39 40.8 43.1 40.2 38 39 40 40 40 P=0.01 for overall F	B Large drop out
GomezSan chiz, 2003, 2004 Spain 296 at birth; 249 at 18 mo; 238 at 24 mo 37-42 wk gestation	1 rural and 1 urban sites, comparing breastfed > 4 mo, <4 mo, and formula-fed; breastfeeding information was from medical records; (reported as a prospective cohort by authors, but parents were informed of the project and gave consent for the study at 15 months check up)	≤ 4 mo 49% > 4 mo 28% (no distinction between exclusive or partial breastfeedin g) Formula-fed 23%	Bayley was administered at 18 and 24 mo	Parental IQ, social class, parental education, number of siblings	Parental IQ was obtained only for 164 couples; their children had MDI 2.3 points higher than the children whose parents did not take part in IQ testing (P<0.05). Infants were breastfed for a mean of 85.7 days ± SD 76.4 days. Duration of breastfeeding had a correlation with MDI at 18 mo (r=0.42; P<0.001) and at 24 mo (r=0.37; P<0.001). At 24 mo, after multiple linear regression adjusting for parental IQ, difference between formula and breastfeeding ≤ 4 mo no longer significant; difference of 4.3 points remained significant when comparing breastfeeding > 4 mo with ≤ 4 mo.	B Bayley at young age

Table 9. Continued

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment tools for cognitive development	Confounders adjusted for	Outcomes	Quality and limitations
Angelsen, 2001 Norway and Sweden Enrolled 521 Followup 291 (5 yr) Term Excluded congenital malformati ons	Prospective cohort; but duration of breastfeeding retrospectively recorded at 13 mo	< 3 mo 17%; 3-6 mo 21%; ≥ 6 mo 62%	Bayley MDI at 13 mo Wechsler Preschool and Primary Scales of Intelligence (WPPSI-R) at 5 y/o	Maternal IQ, age, education, smoking	Bayley MDI at 13 mo: 117.7 (SD11.7) in ≥ 6 mo breastfeeding compared to 109.9 (SD13.1) in < 3 mo (P<0.001). There was a linear increase in MDI plotted against breastfeeding duration (P<0.001). Maternal intelligence (Raven test score) was also related to duration of breastfeeding. Adjustment for differences in maternal intelligence reduced the OR of having a low MDI among children who were breast fed for <3 mo to 1.6 (95%CI 1.1- 2.3) from 3.2 (95%CI 1.7-5.9). Total IQ at 5 y/o: 111 (SD14.3) in ≥ 6 mo breastfeeding compared to 103.6 (SD14.6) in < 3 mo (P<0.001). Adjustment for differences in maternal intelligence reduced the OR of having a low IQ score among children who were breastfed for <3 mo to 1.5 (95%CI 1.0-2.1) from 2.8 (95%CI 1.4-5.3).	B Large drop out
Agostoni, 2001 Italy Enrolled 95 Followup 44 Term infants	Bayley results at 1 yr were compared in breastfeeding ≥6 mo with 3-6 mo	≥ 3 mo 46% 6 mo 31% 9 mo 18% 12 mo 11%	Bayley MDI at 1 y/o	Parity, maternal education, age, and smoking habits	Unadjusted Bayley MDI in 29 subjects breastfed > 6 mo = 94, in 15 subjects breastfed 3-6 mo = 92.1; After adjustment, Bayley MDI in 29 subjects breastfed >6 mo compared to 15 subjects breastfed 3-6 mo showed a 2.0 point advantage (95% CI -3.2, 7.3).	C Small sample size, Bayley at young age

BMI, body mass index; MDI, Mental Development Index; NLSY79, national longitudinal survey of youth 1979; PIAT, Peabody individual achievement test; PPVT-R, Peabody Picture Vocabulary Test; SR, systematic review; MA, meta-analysis; SES, socioeconomic status; WPPSI-R, Wechsler Preschool and Primary Scales of Intelligence

Relationship between Obesity and Breastfeeding

Background

The prevalence of overweight or obesity among children and adolescents has rapidly increased in the past two decades. Results from the 1999-2002 National Health and Nutrition Examination Survey (NHANES) showed that an estimated 16 percent of children and adolescents ages 6-19 years were overweight. This represents a 45 percent increase from the overweight estimate of 11 percent obtained from NHANES III (1988-94). (www.cdc.gov/nchs/products/pubs/pubd/hestats/overwght99.htm) It is increasingly recognized that nutrition in early life may have long-term physiologic effects. Relationships between the types of postnatal feeding and the subsequent development of fat and fat-free mass are quite complex and are dependent on multiple factors including differences in food composition (human milk versus formula), food delivery (breast versus bottle), food “lifestyle” (breastfeeding versus formula feeding) and food behavior (self-regulation and feeding on demand versus set schedules of feeding of predetermined amounts).⁷³ It is known that infants fed breast milk differ in their growth kinetics from formula-fed infants. Formula-fed infants demonstrate higher weight and length gains compared with breastfed infants. A systematic review of 19 studies in developed countries concluded that by the age of 12 months, the cumulative difference in body weight amounted to approximately 400 g less in infants breastfed for 9 months compared with formula-infants, and as much as 600-650 g less in infants breastfed for 12 months compared with formula-fed infants.⁷⁴ Differences in feeding behavior and mother-child interaction between breast- and formula-fed infants may account for some of the differences reported. For instance, breastfed infants showed a different suckling pattern, and appeared to have greater degree of control on meal sizes and feeding intervals than infants who were formula-fed.⁷⁵ Diet-related differences in the circulating levels of biochemical markers (such as leptin, ghrelin, insulin-like growth factors, and other compounds) implicated in energy metabolism during infancy might explain some of the anthropometric and behavioral differences between breastfed and formula-fed infants. These observed differences may have potential long-term consequences.⁷³

Commonly considered confounders in the studies of relationship between obesity or overweight and breastfeeding were birth weight, parental overweight, parental smoking, dietary factors, physical activity, socioeconomic status (SES), age, sex, birth order, and number of siblings.

Additional Methodological Comments

We identified three systematic reviews and meta-analyses that examined the relationship between breastfeeding and childhood obesity or obesity across all ages.⁷⁶⁻⁷⁸ Although the outcomes of interest were similar among these systematic reviews and meta-analyses, they answered slightly different research questions because of the differences in their study eligibility criteria and analyses. Thus, we have summarized and discussed these systematic reviews and meta-analyses separately.

Published Systematic Reviews/Meta-Analyses (Table 10)

Arenz 2004 was a meta-analysis of studies from 1966 to December 2003 that examined the relationship between breastfeeding and childhood obesity in children at least one year of age. Inclusion criteria for the meta-analysis were: obesity defined by a body mass index (BMI) greater than 90th, 95th or 97th percentile; adjustment for at least three potential confounding or interacting factors; reported either odds ratio or relative risk; and last followup between 5 and 18 years of age. Nine of 28 studies reviewed met the eligibility criteria for meta-analysis. There were two prospective cohort and seven cross-sectional studies totaling more than 69,000 children from developed countries. The meta-analysis used both fixed- and random-effects models and pooled crude and adjusted odds ratios from the individual studies. Definitions of breastfeeding and comparative feedings were heterogeneous across studies. Sensitivity analyses were performed to assess for heterogeneity. The factors analyzed were cohort study or cross-sectional study, different definitions of breastfeeding, different definitions of obesity, different age groups and number of potential confounders considered for adjustment. The methodological quality of this meta-analysis was grade A.

The pooled crude odds ratio for breastfeeding and obesity defined as a BMI > 90th, 95th or 97th percentile could be calculated for six of the nine studies included. The odds ratio was 0.67 (95% CI 0.62 - 0.73). The adjusted odds ratio for the nine studies was 0.78 (95% CI 0.71 - 0.85) for both the fixed and random-effects models, suggesting that there was no heterogeneity between the studies. Sensitivity analyses showed that the protective effect of breastfeeding was more pronounced in studies with adjustment for less than seven potential confounding factors compared with adjustment for seven or more potential confounding factors (adjusted OR 0.69 vs. 0.78, respectively). Other criteria (e.g., cohort study or cross-sectional study, different definitions of breastfeeding, different definitions of obesity, different age-groups) did not affect the summary estimates significantly. For example, the pooled adjusted odds ratio of obesity was 0.76 (95%CI 0.67-0.86) in studies comparing ever breastfeeding to never breastfeeding, versus 0.74 (95%CI 0.64-0.85) in studies that used other definitions of breastfeeding and comparative feedings.

Eight studies analyzed the relationship between breastfeeding duration and the risk of overweight or obesity in later childhood. Exclusivity of breastfeeding was not reported. Four studies reported an inverse association of breastfeeding duration and the prevalence of obesity both in the crude and the adjusted estimates. One of the studies lost statistical significance after adjustment. Three studies found no significant effect of duration of breastfeeding on obesity.

Harder 2005 was a meta-analysis of 17 qualifying studies published from 1966 to December 2003. A total of 120,831 subjects (66 to 32,200 subjects per study) from developed countries were included. Eligibility criteria included any original report comparing breastfed subjects with exclusively formula-fed subjects at any age, the reports must have either reported odds ratio or contained data for the calculation of odds ratio for the risk of overweight or obesity in relationship to the feeding history, and the duration of breastfeeding must have been reported. All definitions of overweight or obesity were included. Three different meta-analytic techniques that specifically required the use of crude odds ratios and 95% confidence intervals were employed. Because of suboptimal consideration for potential confounding, we rated the methodological quality of this meta-analyses grade B.

Fourteen studies provided data for more than one category of duration of breastfeeding, leading to 52 estimates included in the meta-regression analysis. In the analysis, duration of

breastfeeding was significantly negatively related to the risk of overweight (regression coefficient: 0.94, 95%CI 0.89 - 0.98). Categorical analysis showed that from 1 month of breastfeeding onward (the reference group), the risk of subsequent overweight continued to decrease, reaching a plateau of more than 30 percent risk reduction at 9 months of breastfeeding. Using the “pool-first method” (that is to calculate a study-specific regression coefficient and corresponding 95 percent confidence interval for each study using a log-linear model and then pooled all studies with a random effects model) to quantify the dose-response relationship, the results showed that each month of breastfeeding was associated with a four percent decrease in risk of overweight per month of breastfeeding exposure (OR 0.96/month of breastfeeding, 95%CI 0.94 - 0.98). A fixed effect model reported a similar pooled odds ratio (OR 0.96/month of breastfeeding, 95%CI 0.95 - 0.98). The age at examination had little influence on the magnitude of the effect of duration of breastfeeding on the risk of overweight. The pooled odds ratio from five studies investigating subjects up to 5 years of age was 0.97 (95%CI 0.94 - 0.99); while for six studies on subjects 6 or more years of age, it was 0.96 (95%CI 0.93 - 0.99).

The effect of the duration of exclusive breastfeeding was analyzed in two studies. The pooled odds ratio for the risk of overweight per month of exclusive breastfeeding was 0.94 (95%CI 0.89 - 0.99, random effects model).

Subgroup analyses showed that the different definitions of overweight influenced the estimate of odds ratio only slightly. In eight studies that used BMI to define overweight, the pooled odds ratio was 0.96 (95%CI 0.94 - 0.98); while in three studies that used another measures (e.g., percentile of weight for length, or weight for age) to define overweight or obesity, the odds ratio was 0.93 (95%CI 0.87 - 0.99).

Lastly, Owen 2005 was a systematic review of 61 observational studies from 1966 to September 2003 that examined the effects of infant feeding on a measure of adiposity (quantitatively or narratively) in later life. Twenty-eight studies (totaling 298,900 subjects) that provided 29 unadjusted odds ratios relating the initial infant feeding method and obesity were included in a meta-analysis. A fixed effect model was used. Meta-regression and sensitivity analyses were used to examine the influence of various factors defined a priori, including the effects of adjustment for factors such as parental body size (mostly BMI), SES, and maternal smoking. Because of suboptimal consideration for potential confounding, we rated the methodological quality of this systematic review and meta-analyses grade B.

Twenty-eight of 29 estimates related breastfeeding to a lower risk of obesity in later life. Four estimates were for infants, 23 for children, and two for adults. There was evidence of marked heterogeneity among studies ($P < 0.001$). In a fixed-effect meta-analysis, breastfed subjects were less likely to be defined as obese than were formula-fed subjects (OR 0.87, 95% CI 0.85 - 0.89). In six studies, it was possible to examine the effect of adjustment for the following potential confounders: SES (based on parental education in two studies), parental BMI, and current maternal smoking or maternal smoking in early life. The pooled odds ratio in these studies changed from 0.86 (95% CI 0.81 - 0.91) before adjustment to 0.93 (95% CI 0.88 - 0.99) after combined adjustment. The effect of adjustment for birth weight (based on either actual birth weight or prevalence of low birth weight) was examined in 10 studies; this had no appreciable effect on the odds ratios.

There was no clear evidence that the protective effect of breastfeeding altered with increasing age of outcome assessment. Odds ratios of 0.50 (95% CI 0.26 - 0.94) for infants, 0.90 (95% CI 0.87 - 0.92) for young children, 0.66 (95% CI 0.60 - 0.72) for older children, and 0.80 (95% CI 0.71 - 0.91) for adults were observed (test for trend, $P = .85$, adjusted for study size; $P = .99$ with

the exclusion of infants). The protective effect of breastfeeding on obesity was stronger and more homogeneous among four studies in which initial feeding groups were exclusive (OR 0.76; 95% CI 0.70 - 0.83; test for heterogeneity between estimates, $P = .143$), compared with all other studies. In 14 studies that provided data on breastfeeding duration, the protective effect of breastfeeding over formula feeding was greater among subjects breastfed for at least 2 months (OR 0.81, 95% CI 0.77 - 0.84), compared with those breastfed for any duration (OR 0.89, 95% CI 0.86 - 0.91). In six studies, it was possible to examine the effect of adjustment for the following potentially important confounders: socioeconomic status, parental BMI, and current maternal smoking or maternal smoking in early life. The pooled odds ratio in these studies was reduced from 0.86 (95% CI: 0.81– 0.91) before adjustment to 0.93 (95% CI: 0.88–0.99) after adjustment.

Thirty-three studies totaling 12,505 subjects explored the relationship between breastfeeding and obesity even though they did not provide odds ratio. However, they provided 35 reports of directions of association; of these, breastfeeding was unrelated to the risk of obesity in 33, related to a reduced risk in one, and related to an increased risk in another. Studies that did not provide odds ratios were much less likely to report that breastfeeding was associated with a reduced risk of obesity, compared with studies that did provide odds ratios (1 of 35 studies and 18 of 29 studies, respectively; $P < .001$).

Conclusion

Findings from three systematic reviews and meta-analyses of good and moderate methodological quality suggest that a history of breastfeeding is associated with a reduction in the risk of obesity in later life. The pooled adjusted odds ratio of overweight/obesity comparing ever breastfeeders to never breastfeeders was 0.76 (95%CI 0.67-0.86) and 0.93 (95%CI: 0.88– 0.99) in Arenz 2004 and Owen 2006 meta-analysis, respectively. In Harder 2005 meta-analysis, duration of breastfeeding was significantly negatively related to the unadjusted risk of overweight (regression coefficient: 0.94, 95%CI 0.89 - 0.98), and each month of breastfeeding was found to be associated with a four percent decrease in risk (unadjusted OR 0.96/month of breastfeeding, 95%CI 0.94 - 0.98). However, the results from Harder 2005 meta-analysis employed techniques that required the use of crude odds ratios from the primary studies for its summary estimates. Therefore, those estimates may not be accurate because potential confounders could not be accounted for in the analysis. As demonstrated in the sensitivity analyses in both Arenz 2004 and Owen 2005, the magnitude of effects was reduced when more confounders were adjusted for in the analyses. The observed association between breastfeeding and a reduced risk of obesity could also reflect selective reporting and/or publication bias. The exclusivity of breastfeeding was not described in the majority of the studies.

Table 10. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and overweight or obesity

Author year Population	Study description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitations
Arenz 2004 Children and adolescents. One study included some adult subjects. N=69,000	MA of 7 cross-sectional and 2 prospective cohort studies in developed countries	Never BF or partly BF < 3 months vs. BF ≥ 3 month; mostly or only BF vs. mostly or only formula feeding in the first 6 months; BF never vs. ever; BF never vs. > 6 months, BF groups: <1 week, 1 week-1 months, 2-3 months, 4-6 months, 7-9 month, > 9 months (exclusivity of BF not reported)	Studies with > 3 the following relevant confounding factors birth weight, parental overweight, parental smoking, dietary factors, physical activity and SES were included. Other confounders in the included studies: age, sex, diet and weight concerns, Tanner stadium, birth order, race, introduction of solid foods, number of siblings	- The pooled adjusted OR for breastfeeding and obesity defined as BMI >90th, 95th or 97th percentile calculated for nine studies was 0.78 (95%CI 0.71-0.85) for both fixed and random-effects model. - Protective effect of breastfeeding was more pronounced in studies with adjustment for less than 7 (but more than 3) potential confounding factors compared to adjustment for seven or more potential confounding factors (OR 0.69 vs. 0.78 respectively). - Pooled adjusted OR of obesity was 0.76 (95%CI 0.67-0.86) compared ever breastfeeding to never breastfeeding, while it was 0.74 (95%CI 0.64-0.85) in studies using other definitions of breastfeeding.	A ^a
Harder 2005 Children and adolescents. Two studies included adult subjects. N=120,831 (ranged from 66 to 32,200)	MA of 16 cohort or cross-sectional studies and 1 case-control study in developed countries	Median duration of breastfeeding categories: < 1 month (reference), 1-3 month, 4-6 months, 7-9 months, > 9month. To studies that provided data for more than two categories of duration of breastfeeding, "pool-first method" was used to quantify the dose-response relation (per month of breastfeeding).	Age, sex, birth weight, SES, Tanner stage, physical activity, eating habits, concerns to gain weight, birth order, dietary intakes, maternal BMI, maternal smoking	- In the weighted meta-regression (52 estimates from 14 gave data for more than one category of duration of breastfeeding), duration of breastfeeding was significantly negatively related to risk of overweight (regression coefficient: 0.94, 95%CI 0.89-0.98). - From 1 month of breastfeeding onward, the risk of subsequent overweight continuously decreased up to a reduction of more than 30%, reaching a plateau at 9 months of breastfeeding. - Each month of breastfeeding was found to be associated with a 4% decrease in risk (OR: 0.96/month of breastfeeding, 95%CI 0.94-0.98).	B Only unadjusted ORs were combined, although some primary studies had adjustments for potential confounding

Table 10. Continued

Author year Population	Study description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitations
Owen 2005 28 studies provided 29 unadjusted odds ratios relating the initial infant feeding method and obesity. Four observations were for infants, 23 for children, and 2 for adults. N=298,900 in meta-analyses N=12,505 in the studies that did not provide an estimate of relative risk	SR of 61 studies that compared a measure of obesity (quantitatively or narratively) among breastfed and formula-fed subjects; with meta-analyses of 28 studies reported sufficient data	Breastfed vs. formula-fed	Parental body size (mostly BMI), socioeconomic status, maternal smoking, physical activity, weight gain during pregnancy, introduction of solid foods, birth order, dietary intakes, maternal BMI, maternal smoking, and parental obesity/overweight	- In a fixed-effects model including all 28 studies, breastfed subjects were less likely to be defined as obese than were formula-fed subjects (OR: 0.87; 95% CI: 0.85–0.89). There was evidence of marked heterogeneity among studies ($p < 0.001$). - Data from 6 studies, the pooled odds ratio was reduced from 0.86 (95% CI: 0.81–0.91) before adjustment for confounders to 0.93 (95% CI: 0.88–0.99) after adjustment. - In 14 studies with information on breastfeeding duration, the protective effect of breastfeeding over formula feeding was greater among subjects breastfed for ≥ 2 months (odds ratio: 0.81; 95% CI: 0.77–0.84), compared with those breastfed for any duration (odds ratio: 0.89; 95% CI: 0.86–0.91) in the same studies. 35 studies (12,505 participants), reported only directions of association; of these, breastfeeding was unrelated to risk of obesity in 33, related to a reduced risk in 1, and related to an increased risk in 1.	B Only unadjusted ORs were combined, although some primary studies had adjustments for potential confounding

BF, Breastfed; BMI, Body Mass Index; SR, systematic review; MA, meta-analysis; MAs, meta-analyses; SES, socioeconomic status

^a Results from overall MA was inappropriate due to heterogeneous definitions of breastfeeding and comparators across studies. However, subgroup analyses on different breastfeeding definitions were performed.

Relationship between the Risk of Cardiovascular Diseases and Breastfeeding

Background

Abnormal levels of total cholesterol, low-density lipoprotein (LDL) cholesterol, and blood pressure in adults are major risk factors for cardiovascular disease (CVD). Observational studies of large cohorts of men and women have consistently reported that serum cholesterol level >200 mg/dL was associated with an increased risk of all-cause and CVD mortality.^{79,80} Each increment of 20 mm Hg of systolic blood pressure and 10 mm Hg of diastolic blood pressure doubles the risk for CVD.⁸¹ Diet modification, weight reduction, and pharmacotherapy can reduce the risk of CVD. We identified one meta-analysis that evaluated the relationship of breastfeeding during infancy with cholesterol levels in adolescents and adults,⁸² two meta-analyses that evaluated the relationship of breastfeeding with adult blood pressure,^{83,84} and one systematic review and meta-analysis that examined the relationship between breastfeeding and cardiovascular disease mortality in later life.⁸⁵

Cholesterols

Published Systematic Reviews/Meta-Analyses (Table 11)

One meta-analysis of 37 cohort and cross-sectional studies evaluated the effect of breastfeeding on total and LDL cholesterol levels among infants, adolescents, and adults.⁸² All primary studies published in English language from both developed and developing countries that reported estimates of a mean difference and standard error in cholesterol levels between breastfed and formula-fed infants were included in the meta-analysis. The meta-analysis utilized a random effects model. A total of 5,829 breastfed and 4,852 formula-fed subjects were evaluated. Data from the included primary studies were categorized into three age strata: infancy (< 1 year of age), children and adolescents (1 to 16 years), and adults (17 years to 65 years). Outcomes included total and LDL cholesterol levels. No information was provided on the timing of the sample collection in relation to fasting or not fasting. Of the 37 studies, there were 26 outcomes (total or LDL cholesterol levels) in infants, 17 in children and adolescents, and nine in adults. The analysis combined data from a broad age category for the adult participants. It was unclear if there were adjustments for potential confounders such as body mass index, height, and socioeconomic status for the data on cholesterol. The methodological quality of the meta-analysis was rated grade C.

In 25 of 26 observations, infants who were breastfed reported higher mean total cholesterol levels compared with infants who were formula-fed. The overall mean difference was +24.75 mg/dL (95% CI 18.95 to 30.55). There was a statistically significant heterogeneity across the studies. The meta-analysis did not find an association between total cholesterol level and age or gender. There were only seven reported observations on LDL cholesterol levels in infants, six of which reported higher mean levels of LDL cholesterol in breastfed infants compared with formula-fed infants. The mean difference was +22 mg/dL (95% CI 15.47 to 29.0 mg/dL). There was no statistical heterogeneity.

In 16 of 17 observations, the mean total cholesterol levels in children or adolescents who were breastfed in their infancy were similar to those who were formula-fed. The overall mean difference was 0.0 mg/dL (95% CI -2.7 to 2.7 mg/dL). A statistically significant heterogeneity was observed across studies. There was no association between total cholesterol level and age or gender. There were only four observations of LDL cholesterol levels in children or adolescents. The mean levels of LDL cholesterol in children or adolescents who were breastfed in their infancy were similar to those who were formula-fed. The mean difference was +0.39 mg/dL (95% CI -2.7 to 3.09 mg/dL). There was no statistical heterogeneity between the studies.

The mean age of adults evaluated in the primary studies ranged from 17 to 64 years. Lower mean total cholesterol levels in adults who were breastfed in their infancy compared with those who were formula-fed in their infancy were reported in seven of nine observations. The overall mean difference was -6.96 mg/dL (95% CI -2.32 to -11.6 mg/dL). There was no statistically significant heterogeneity between the studies. There were only four observations of LDL cholesterol reported in adults. Adults who were breastfed in their infancy had lower mean LDL cholesterol levels compared with those who were formula-fed. The mean difference was -7.7 mg/dL (95% CI -3.09 to -12.37 mg/dL). There was no statistical heterogeneity between the studies.

Conclusion

Results from the meta-analysis of cohort and case-control studies reported that there was a reduction in total and LDL cholesterol levels in adults who were breastfed during infancy compared with those who were formula-fed. While higher serum lipid levels were observed in infancy, the meta-analysis found that breastfeeding was associated with a reduction in serum lipid level in adult life. The significance of higher serum lipid levels observed in infancy is unclear and studies have neither shown benefit nor harm from such high levels.⁸⁶ These findings were based on data from adults with a wide age range. The analysis did not segregate the data according to gender. Potential confounders were not explicitly analyzed. Detailed information (e.g., fasting or non-fasting) on the collection of specimen for cholesterol testing was not included. The methodological quality of the meta-analysis was rated grade C. Because of the poor methodological quality of the meta-analysis, we find that the conclusions drawn by the authors were suspect. We conclude that the relationship between breastfeeding and adult cholesterol levels cannot be correctly characterized at this time.

Table 11. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and total cholesterol and low-density lipoprotein cholesterol

Author year	Number of subjects	Population	Confounders considered	Study description	Intervention / Comparator	Results		Quality of SR/MA
		3 age strata:				Combined mean difference (95%CI)		
						Total cholesterol (mg/dL)	LDL cholesterol (mg/dL)	
Owen 2002	5,829 vs. 4,852	Age <1 yr	None	MA of 37 cohort and cross-sectional studies	Breastfeeding vs. formula feeding in infancy	+24.75 (18.95 to 30.55)	+22.0 (15.47 to 29.0)	C
		1 to 16 yr				0.0 (– 2.7 to 2.7)	+0.39 (–2.7 to 3.09)	
		≥17 to 65 yr				–6.96 (2.32 to 11.6)	–7.7 (3.09 to 12.37)	

LDL, low-density lipoprotein; SR, systematic review; MA, meta-analysis

Blood Pressure

Published Systematic Reviews/Meta-Analyses (Table 12)

Two meta-analyses evaluated a total of 26 studies of various designs for the effect of breastfeeding and formula feeding in infancy on systolic and diastolic blood pressure levels in adult life. The primary studies included in these meta-analyses were conducted in developed and developing countries. Outcomes assessed were differences in systolic and diastolic blood pressures in adulthood. Both meta-analyses examined potential confounders in the studies and also explored the possibility of publication bias. Potential confounders considered in the studies were age, gender, race, height, and body mass index. Among the 26 primary studies evaluated, 13 studies were common to both meta-analyses. Both meta-analyses included at least one study that included preterm infants. Duration and exclusivity of breastfeeding or formula feeding were heterogeneous in the studies. The methodological quality of both meta-analyses were rated grade B.

The meta-analysis by Martin 2005 included 15 studies with 17 observations published until 2004 that evaluated a total of 17,503 eligible subjects.⁸³ In these studies, blood pressure was measured in subjects whose age ranged from 1 to 60 years. The meta-analysis utilized a random effects model. Meta-regression was performed to evaluate the effect of heterogeneity in study size, age at measurement of blood pressure (<10 years, 11-45 years, >45 years), maternal recall, exclusivity of the feeds, methods of blood pressure measurement, and other variables on the summary estimates. Mean systolic blood pressures were reduced by 1.4 mm Hg (95% CI 0.6 to 2.2) in adulthood among subjects who were breastfed in infancy compared with those who were formula-fed. There was a statistically significant heterogeneity across studies for this outcome. Mean diastolic blood pressures were reduced by 0.5 mm Hg (95% CI 0.04 to 0.9) in adulthood among subjects who were breastfed in infancy compared with those who were formula-fed, with no heterogeneity for this outcome.

Owen 2003 included 24 studies published until 2003 with 26 observations for systolic blood pressure and 23 observations for diastolic blood pressure.⁸⁴ The meta-analysis evaluated 8,471 subjects who were breastfed in infancy and 11,292 who were formula-fed. In these studies, blood pressure was measured in subjects whose age ranged from 1 to 71 years. The meta-analyses

utilized a random effects model. Subgroup analyses evaluated the effect of heterogeneity in study size, age of assessment of outcomes, and year of birth. Mean systolic blood pressures were reduced by 1.10 mm Hg (95% CI 0.42 to 1.78) in adulthood among subjects who were breastfed in infancy compared with those who were formula-fed. When stratified by age groups (≤ 1 year, >1 to 16 years, ≥ 17 years), a similar association was observed among age groups that ranged from more than 1 year to 16 years. There was a statistically significant heterogeneity for this outcome. Mean diastolic blood pressures were reduced by 0.36 mm Hg (95% CI -0.08 to 0.79) among subjects who were breastfed compared with subjects who were formula-fed. There was no statistical heterogeneity for this outcome.

Both meta-analyses observed smaller association of breastfeeding on systolic blood pressure in large studies (>300 and/or >1000 participants) compared with smaller studies. Both meta-analyses attributed this observation to publication bias. Similar effect size reduction of breastfeeding on diastolic blood pressure was not observed. In studies that adjusted for potential confounders, results remained similar before and after adjustment for potential confounders.

Conclusion

Results from both meta-analyses concluded that there was a small reduction in systolic blood pressures among adults who were breastfed in their infancy compared with those who were formula-fed. The association weakened after stratification by study size, suggesting the possibility of bias. Although both analyses had moderate methodological quality and reported similar findings, the authors had different appraisals of the public health importance of the small reduction in systolic blood pressure. In conclusion, there is an association between a history of breastfeeding during infancy and a small reduction in adult blood pressure, but the clinical or public health implication of this finding is unclear.

Table 12. Summary table for systematic reviews/meta-analyses on the relationship between breastfeeding in infancy and blood pressure levels later in life

Author year Population	Confounders considered	Study description	Intervention / Comparator	Results		Quality of SR/MA Limitations
Martin 2005 N=17,503 Age > 1to 60 yr	Considered but not pooled: age, gender, race, height, and body mass index	Meta-analysis of 15 studies (Observations within 2RCTs, 8 prospective cohorts, 1 retrospective cohort, and 4 cross-sectional studies)	Breastfeeding vs. formula feeding in infancy	Combined mean difference (95%CI)		B No pooled adjusted estimate
				SBP mmHg	DBP mmHg	
				-1.4 (-2.2, -0.6) p=0.001	-0.5 (-0.9, -0.04) p=0.03	
Owen 2003 N=8471 vs. 11,292 Age > 1to 71 yr		Meta-analysis of 24 studies (Observations within 1RCT, 12 cross-sectional, and 11 cohort studies)	Breastfeeding vs. formula feeding in infancy	Combined mean difference (95%CI)		B No pooled adjusted estimate
				SBP mmHg	DBP mmHg	
				-1.10 (-1.8, -0.4)	- 0.36 (-0.79, 0.08)	

SBP, systolic blood pressure; DBP, diastolic blood pressure; SR, systematic review; MA, meta-analysis

Cardiovascular Disease Mortality

Published Systematic Review/Meta-Analysis (Table 13)

We identified one systematic review and meta-analysis that examined the relationship between breastfeeding and cardiovascular disease (CVD) mortality in later life.⁸⁵ Articles were included in the systematic review if breastfed infants were compared with bottle-fed infants, if the outcome was cardiovascular disease or ischemic heart disease mortality, and if estimates of the association between having been breastfed and cardiovascular disease or ischemic heart disease mortality could be obtained from the paper or after correspondence with the authors. A total of four historical cohort studies from developed countries were identified, involving 25,166 subjects at baseline and 10,785 subjects at followup. The studies were not graded for their methodological quality. The meta-analysis used a random-effects model. Heterogeneity across studies was assessed. The methodological quality of this systematic review and meta-analysis was rated grade B due to incomplete consideration of the heterogeneity across studies in the meta-analyses.

The four historical cohorts included in the systematic review and meta-analysis were Wingard cohort, 1,373 birth children in California (85 percent in followup); Hertfordshire cohort, 5,908 women and 10,374 men born in Hertfordshire (43 percent in followup); Boyd Orr cohort, 4,999 men and women from a survey of diet and health in pre-war Britain (71 percent in followup); and Caerphilly cohort, 2,512 middle-aged men living in Caerphilly, South Wales (63 percent in followup). Subjects from these four cohorts were born between 1904 and 1939. Potential confounders considered in the association between the risk of CVD mortality and breastfeeding in the four studies were age, birth weight, infant health, socioeconomic status, and/or birth order.

Random-effects model showed little difference in all cause mortality between breast- and bottle-fed subjects (pooled rate ratio = 1.01; 95% CI: 0.91 - 1.13, P=0.8), and there was little evidence of heterogeneity. Five observations from three studies suggested little or no association between breastfeeding and cardiovascular disease mortality in both males and females, and one suggested a possible adverse effect (Caerphilly cohort). In random effects meta-analysis, CVD mortality was similar in breastfed versus bottle-fed subjects (pooled rate ratio = 1.06; 95% CI: 0.94 – 1.20), and there was no statistical evidence of between-study heterogeneity.

Ischemic heart disease mortality was 6 percent lower among males who had been breastfed in the Hertfordshire cohort, but 56 percent higher among breastfed females. This result was in agreement with point estimates from the Boyd Orr cohort, suggesting that ischemic heart disease mortality was 10 percent lower among males who had been breastfed, but 40 percent higher among breastfed females (although there was little statistical evidence of interaction: P = 0.2). In Caerphilly, however, ischemic heart disease mortality was 73 percent higher among breastfed males. In a random effects meta-analysis (pooled rate ratio = 1.19; 95% CI 0.89 - 1.58, P = 0.3), and there was evidence of heterogeneity.

Similar analyses were also performed to examine the association between prolonged breastfeeding (> 1 year duration) and the risk of all-cause, CVD, and ischemic heart disease mortality in later life. There was little evidence that prolonged breastfeeding was associated with all-cause mortality (pooled rate ratio: 0.94; 95% CI 0.71 – 1.24), although there was moderate statistical evidence of heterogeneity. There was some evidence that prolonged breastfeeding was associated with a 16 percent increase (95% CI 0.99 – 1.36; P = 0.06) in CVD mortality, and no

evidence of inconsistency in estimates. There was little evidence that prolonged breastfeeding was associated with ischemic heart disease mortality (rate ratio: 1.08; 95% CI 0.88 – 1.31; P = 0.5) and there was no heterogeneity.

Conclusion

The authors concluded that the data reviewed did not provide evidence that breastfeeding was related to all-cause or CVD mortality. The confidence limits around the point estimates and the observed between-study heterogeneity for associations between breastfeeding and ischemic heart disease, however, do not rule out important beneficial or adverse cardiovascular effects of breastfeeding.

There were some possible sources of bias and limitations in the studies reviewed. Two of the four studies had followup rate of less than 70 percent of the original population; therefore, selection bias cannot be ruled out. Recall bias was possible in the three studies where breastfeeding data were collected retrospectively. As confounding and biases may have distorted results from individual studies, the statistical combination of estimates into a single rate ratio needs to be interpreted with caution. All four studies in the meta-analyses were historical cohorts (born between 1904 and 1939). Given the statistical heterogeneity across studies, combining study results might not be appropriate. For the outcome of ischemic heart disease mortality, it may not be appropriate to combine results for men and women into a single analysis because of apparent effect modification by gender.

Because of the above limitations, no definitive conclusions can be drawn regarding the relationship between breastfeeding and CVD mortality. Further investigation is warranted.

Table 13. Summary of systematic review/meta-analysis on the relationship between breastfeeding and long-term cardiovascular disease mortality

Author year Population	Study description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitations
Martin 2004	MA of 4 historical cohort studies in developed countries	Any or exclusively breastfed vs. bottle- fed; prolonged (>1 yr) breastfed vs. bottle-fed	Age, birth weight, infant health, socioeconomic status, number of siblings, and/or birth order	All four studies were historical cohorts born between 1904 and 1939. Five observations from three studies suggested little or no association between breastfeeding and CVD mortality in both males and females (pooled rate ratio: 1.06; 95% CI: 0.94–1.20), and one suggested a possible adverse effect (Caerphilly cohort). There was no statistical evidence of between-study heterogeneity. IHD mortality was 6% lower amongst males who had been breastfed in the Hertfordshire cohort, but 56% higher amongst breastfed females. This result is in line with point estimates from the Boyd Orr cohort suggesting that IHD mortality was 10% lower amongst males who had been breastfed, but 40% higher amongst breastfed females (although there was little statistical evidence of interaction: $p=0.2$). In Caerphilly, however, IHD mortality was 73% higher amongst breastfed males. In a random effects meta-analysis (pooled rate ratio: 1.19; 95% CI 0.89–1.58, $p=0.3$), there was evidence of heterogeneity. There was weak evidence that prolonged breastfeeding was associated with a 16% increase (95% CI: 0.99– 1.36; $p=0.06$) in CVD mortality, and no evidence of inconsistency in estimates. There was little evidence that prolonged breastfeeding was associated with IHD mortality (rate ratio: 1.08; 95% CI: 0.88–1.31; $p=0.5$) and there was no heterogeneity.	B Men and women should be considered separately because of apparent effect modification by gender
Wingard (1994) cohort: 1373 birth children in California (85% in follow-up)					
Hertfordshire cohort (update to 1999): 5908 women and 10374 men born in 11 of 12 districts in Hertfordshire (43% in followup)					
Boyd Orr (2003): 4999 men and women from a survey of diet and health in pre-war Britain (71% in followup)					
Caerphilly (2003): 2512 middle-aged men living in Caerphilly, South Wales (63% in followup)					
Total N=25,166 at baseline; 13,302 at followup					

IHD, Ischemic heart disease; CVD, cardiovascular disease; SR, systematic review; MA, meta-analysis

Relationship between Type 1 Diabetes and Breastfeeding

Background

Type 1 diabetes results from destruction of the insulin-producing β cells of the pancreatic islets. Various exogenous triggers, such as certain dietary factors and viruses, are thought to induce the immune-mediated process leading to extensive β cell destruction.⁸⁷ Putative mechanisms of protection against type 1 diabetes afforded by breast milk include passive immunity provided by secretory immunoglobulin A antibodies, increased β cell proliferation observed in breastfed compared with formula-fed infants, or delayed exposure to foreign food antigens in exclusively breastfed infants.⁸⁷ In addition, hypotheses have been proposed to explain the putative diabetogenicity of cow milk.⁸⁷ For example, β -lactoglobulin, a cow milk-specific protein, has been implicated as a possible trigger of the immune defect, leading to type 1 diabetes. Therefore, concerns regarding the safety and advisability of feeding cow milk-based products to infants have been raised.⁸⁸ Gersterin 1994 conducted a systematic review of the epidemiological and clinical literature that explored a possible link between cow milk and type 1 diabetes.⁸⁹ Subsequently, Norris and Scott 1996 performed meta-analyses on all published case-control studies from 1966 to 1994 that examined infant diet exposures and type 1 diabetes.⁹⁰ Since this meta-analysis, we identified six case-control studies that reported outcome of type 1 diabetes in relation to breastfeeding during infancy.

Commonly considered confounders in the relationship between type 1 diabetes and breastfeeding were maternal/parental education, maternal age at birth, birth order, household income, race/ethnicity, social class, family history of type 1 diabetes, neonatal illness, and type of delivery.

Published Systematic Reviews/Meta-Analyses (Table 14)

Gersterin 1994 conducted a systematic review on literature that explored a possible link between cow milk and type 1 diabetes. Articles were excluded if they exclusively used surrogate markers for either type 1 diabetes or cow milk exposure. A total of three ecological and time-series studies, 13 case-control studies, one cohort study, and one case series were included. Meta-analysis of all included case-control studies was performed, using a fixed effect model. Both adjusted and unadjusted odds ratios were used, but there was no further analysis or discussion on potential impacts of confounding by combining unadjusted odds ratios. The methodological quality of the systematic review and meta-analysis was rated grade B, due to insufficient consideration for the potential confounding in the primary studies.

In the included time-series and ecological studies, the results showed an inverse association (geographical and temporal) between the rate/prevalence of breastfeeding and the rate/prevalence of type 1 diabetes.

The cohort study was an analysis of two groups of children born in the UK in 1958 and 1970, who were followed for 16 and 10 years, respectively. This study did not find any association between breastfeeding for less than 1 month and the risk of development of type 1 diabetes.

In the 13 case-control studies in which the neonatal feeding history of patients with type 1 diabetes and individually matched non-diabetic control subjects were compared, the results were mixed. A total of 3,708 type 1 diabetes cases and 20,340 non-diabetic controls were included. None of these studies satisfied all six methodological criteria defined by the author. Four of the 13 case-

control studies fulfilled five of six methodological criteria (e.g., inclusion of ≥ 75 percent of eligible diabetic patients, unbiased selection of unrelated nondiabetic control subjects, control subjects derived from the same population as diabetic subjects, an identical means of determination of infant feeding practices in both diabetic and nondiabetic groups, blind determination of early feeding history, and identification of diabetic patients from incident cases). When these four studies were combined, the overall odds ratio (adjusted odds ratios were used) for type 1 diabetes in patients exposed to less than 3 months of breastfeeding was 1.43 (95%CI 1.15 - 1.77; $P=0.3$ for homogeneity). When all 13 included case-control studies were combined, the overall odds ratio (mixed crude and adjusted odds ratios were used) for type 1 diabetes in patients exposed to less than 3 months of breastfeeding was 1.37 (95%CI 1.22-1.53; $P=0.11$ for homogeneity).

In 1996, Norris and Scott performed a meta-analysis of infant diet and type 1 diabetes to examine further the inconsistent results reported in the literature. A total of 17 case-control studies with appropriate data for meta-analysis were included. The analysis included 4,656 type 1 diabetes cases and 16,383 non-diabetic controls. The authors abstracted case and control data for four separate exposures: breastfeeding status (ever/never), total breastfeeding duration, exposure to breast-milk substitutes, and exposure to cow milk-based substitutes. Due to the limitation of the meta-analytic technique, only unadjusted odds ratios were calculated and combined, although some primary studies had adjustments for maternal education, maternal age at birth, birth order, household income, race/ethnicity, and/or social class. Authors performed sensitivity analyses on various characteristics of study methodological quality to explore the impacts of potential biases on the summary odds ratios. The methodological quality of this meta-analysis was rated grade B, due to insufficient consideration for the potential confounding in the primary studies.

The overall odds ratio of all included case-control studies that examined the association between never breastfed and type 1 diabetes was 1.13 (95%CI 1.04 - 1.23). Among these studies, fourteen also examined type 1 diabetes risk by months of breastfeeding duration. The duration categories in the analysis were cumulative, rather than mutually exclusive (i.e., some studies provided data on both 3-month and 6 month breastfeeding data in the same subjects). The summary odds ratios showed consistently elevated risks of type 1 diabetes associated with age at first exposure to any breast milk substitutes before 6 months of age. Since the majority of the studies reported odds ratios using a cutoff of 3 months when examining continuous exposures, this cutoff was used for the meta-analysis. The summary odds ratio for type 1 diabetes in subjects who were breastfed for less than 3 months compared with those who were breastfed for at least 3 months was 1.23 (95%CI 1.12 - 1.35).

Stratified analyses of studies by methodological and study population characteristics were performed to see whether differences in these characteristics might explain the heterogeneity. The characteristics were prevalent versus incident case-control study design, adequate versus inadequate response rates of the cases and controls, the breastfeeding prevalence in the background population, the type 1 diabetes risk in the background population, and retrospective versus concurrent infant diet assessment. All of these factors had differential impacts on the summary odds ratios for the risk of type 1 diabetes associated with infant diet exposures.

Norris and Scott concluded that their meta-analysis showed that the increased risk of type 1 diabetes associated with any of the infant diet exposures was small. According to these authors, interpretation of weak associations (i.e., an odds ratio of less than 2.0) can be problematic, since weak associations can more readily be explained by biases.

Studies Identified after the Published Systematic Reviews/Meta-Analyses/Systematic Review (Table 15)

We only included all studies that examined the outcome of type 1 diabetes in relation to breastfeeding in developed countries. To be consistent with previous meta-analyses, articles were excluded if they only used surrogate markers for type 1 diabetes (e.g., the presence of islet cell antibodies). A total of six case-control studies were identified,⁹¹⁻⁹⁶ The studies included 1,293 patients with type 1 diabetes and 3,262 control subjects. Five studies were conducted in Europe, and one in Taiwan. Four studies were rated grade B in methodological quality; two studies were rated grade C. Commonly considered confounders in these studies were family history of type 1 diabetes, neonatal illness, maternal age at birth, birth order, maternal/parental education, and type of delivery.

In four studies, the definition of type 1 diabetes cases was children with juvenile-onset diabetes (or developing diabetes before 17 years of age). In the other two studies, the cases were registered type 1 diabetes patients who were younger than 30 years of age and children with diabetes who were identified through hospital records. Matched controls were those without type 1 diabetes selected from various sources in the same population as the cases.

Three studies reported odds ratio of type 1 diabetes comparing subjects who were ever breastfed with those who were never breastfed. Two studies found a reduced risk of type 1 diabetes in subjects who were ever breastfed (ORs 0.56⁹⁵ and 0.75⁹¹), while the third study reported an increased risk (OR 2.44⁹⁴).

Three studies compared subjects who were breastfed for more than 3 months or 6 months with those who were never breastfed. Two studies reported a reduced risk of type 1 diabetes in subjects who were breastfed for more than 3 or 6 months (adjusted OR 0.57⁹⁶ and 0.25⁹⁵, respectively). The third study⁹⁴ reported an opposite finding (this was the same study that reported an increased risk with ever breastfed). A reduced risk of type 1 diabetes was found when comparing subjects who were never breastfed with those who were breastfed for more than 6 months (adjusted OR 0.36, 95%CI 0.14 - 0.94). The data also showed a slight increase in the risk of type 1 diabetes with longer duration of breastfeeding (1 month increment). The control subjects in this study were selected from children admitted to the same hospital as the cases, whether this explained the finding was unclear.

One study reported a reduced risk of type 1 diabetes comparing subjects who were initially exclusively breastfed with those who were not (adjusted OR 0.6, 95%CI 0.41 - 0.89).⁹³ Another study found a small, but non-significant increased risk of type 1 diabetes with not breastfeeding at discharge (RR 1.33, 95%CI 0.76 - 2.31).⁹² This study was rated to have poor methodological quality because only univariate analysis was performed and potential confounders were not considered.

Conclusion

Our findings from the six additional case-control studies are similar to the findings from the two meta-analyses. However, the exclusivity of breastfeeding was not addressed in all studies, and the assessment of infant diet was based on long-term recall in five of six studies. We elected not to perform a meta-analysis, because it is unlikely to change the pooled estimates from the previous meta-analyses by adding additional three studies from the updates that compared subjects who were breastfed for more than 3 or 6 months with those who were never breastfed.

Two meta-analyses of moderate methodological quality reported statistically significant odds ratios of 1.23 and 1.43, respectively, for the risk of type 1 diabetes in subjects exposed to less than 3 months compared with more than 3 months of breastfeeding. Since case-control studies are prone to recall biases, Norris and Scott compared the odds ratios in studies relied on long-

term recall to assess infant diet with studies that did not. The results showed that studies using existing infant records to determine breastfeeding initiation and duration failed to show the associations reported in the studies relying on long-term recall for their exposure data. This suggests that subjects with type 1 diabetes were more likely to report shorter duration of breastfeeding than control subjects.

In conclusion, even though there is some evidence to suggest that breastfeeding for more than 3 months is associated with a reduced risk of type 1 diabetes, this evidence must be interpreted with caution because of the possibility of recall biases and suboptimal adjustments for potential confounders in the primary studies.

Table 14. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and type 1 diabetes

Author Year Population	Study description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitations
Norris 1996 Case-control studies in which the neonatal feeding histories of patients with type 1 DM and individually matched non-DM control subjects were compared Cases: 4,656 Controls: 16,383	MA of 17 case-control studies published from 1966 to 1994.	A=breastfeeding status (ever/never) B=total breastfeeding duration C=exposure to breast-milk substitutes D=exposure to cow's milk-based substitutes	Maternal education, maternal age at birth, birth order, household income, race/ethnicity, social class	The summary OR for type 1 DM in patients who were never being breast-fed was 1.13 (95% CI 1.04-1.23). The summary OR for type 1 DM in subjects who were breast-fed for < 3 mo compared with those who were breast-fed for at least 3 mo was 1.23 (95%CI 1.12-1.35). Stratified analyses of studies were performed by prevalent vs. incident case-control study design, adequate vs. inadequate response rates of the cases and controls, the breastfeeding prevalence in the background population, the type 1 DM risk in the background population, and retrospective vs. concurrent infant diet assessment. All of these factors had impacts on the summary ORs for the risk of type 1 DM associated with infant diet exposures.	B Only unadjusted ORs were combined although some primary studies had adjustments for potential confounding
Gerstein 1994 Case-control studies in which the neonatal feeding histories of patients with type 1 DM and individually matched non-DM control subjects were compared Cases: 3,708 Controls: 20,340	A SR review of 3 ecological and time-series studies, 13 case-control studies, one cohort study, and one case series. MA was performed for the case-control studies.	Cow's milk exposure (and therefore non-exclusive breastfeeding) vs. Cow's milk avoidance (and therefore breastfeeding)	ND	Results from the 13 case-control studies were mixed. The combined OR for type 1 DM in patients exposed to < 3 mo of breastfeeding was 1.38 (95% CI 1.22-1.53; p=0.11 for homogeneity). An analysis of 2 cohorts of children born in the UK in 1958 and 1970 followed for 16 and 10 yr, respectively, failed to show any association between breastfeeding for < 1 mo and type 1 DM. None of the 13 case-control studies fulfilled all six methodological criteria. Four case-control studies (31%) fulfilled five out of the six criteria (considered high quality)	B Combine both crude and adj. ORs without performing sensitivity analyses on potential confounding or study quality

MA, meta-analyses; SR, systematic review

Table 15. Summary of case-control studies on the relationship between breastfeeding and type 1 diabetes

Author year Country	Cases (N)	Control (N)	Definition of type 1 DM	Mean Age at Dx (year)	Breastfeeding group	Comparator group	OR* (95% CI)		Potential confounders adjusted	Quality and Limitations
							Crude	Adjust		
EURODIAS 2002 Europe	610	1616	Onset < 15 yr in 1989- 1995	ND	Ever BF	Never BF	0.75 (0.58-0.96)	0.59 (0.35-0.97)	Ht SDS, Wt SDS, maternal age, jaundice, RTI, Vit D suppl, asthma [Sex, age] ^a , maternal age, maternal type 1 DM, preeclampsia, C-section, neonatal illnesses	B No demographic data; No adj. for SES
McKinney 1999 UK	196	325	Onset < 16 yr in 1993- 1994	0-15	Initial exclusive BF	Initial not exclusive BF	0.68 p=0.04	0.60 (0.41-0.89)		B No adj. for SES
Tai 1998 Taiwan	177	193	Registered type 1 DM, age < 30 yr, born in 1984-1993	8.3	Ever BF	Never BF	0.56 (0.40-1.2)		[Age, sex, parental education] ^a , birth order, paternal age, GA, type of delivery, BW, monthly family income	B Inconsistent methods for ascertainment of BF exposure
					BF < 6 mo	Never BF	0.82 (0.47-1.42)	0.84 (0.45-1.59)		
					BF ≥ 6 mo		0.39 (0.16-0.94)	0.25 (0.09-0.69)		
Meloni 1997 Italy	100	100	Onset < 17 yr in 1983- 1994	6 (1-15)	Ever BF	Never BF		2.44 (1.10-7.14)	[Age, sex] ^a , mother's education, number of siblings	B Hospital controls
						BF 3-5 mo		0.85 (0.37-1.92)		
					BF > 6 mo	BF 1-2 mo		2.08 (0.18-5.26)		
						BF 0 mo		2.78 (1.06-7.14)		
					BF as continuous coefficient (1 mo increment)			1.10 (0.99-1.22)		
Visalli 2003 Italy	150	750	Onset < 15 yr, born in 1977- 1989	ND	BF ≥ 3 mo	BF < 3 mo	0.47 (0.31-0.72)	0.57 (0.41-0.74)	[Age] ^a , family history of type 1 DM, infectious disease during pregnancy, eczema	C Inadequate response rate (>20%); No adj. For SES
Jones, 1998 UK	60	458	Dx at discharge, born in 1976-1986	ND	BF at discharge	Not BF at discharge	0.75 (0.43-1.32)		[Age, sex] ^a	C Poor adj. for confounding

Dx, diagnosis; BF, breastfeeding; SR, systematic review; MA, meta-analysis; SDS, standard deviation score; Ht, height; Wt, weight; RTI, respiratory infection; Vit D suppl, Vitamin D supplementation; C-section, delivery by cesarean section; GA, gestational age; BW, birth weight; LTC, long-term recalls; adj, adjustment; SES, socioeconomic status

*Odds ratio of type 1 DM, compared the breastfeeding group to the comparator group (or the reference group), unless noted

^a Matching factors for controls

Relationship between Type 2 Diabetes and Breastfeeding

Background

In 2002, it was estimated that a total of 18.2 million people, or 6.3 percent of the US population carried a diagnosis of diabetes (cdc.gov/diabetes/pubs/estimates.htm#prev). No data are currently available on the prevalence of type 2 diabetes in children and adolescents. The Centers for Disease Control and Prevention (CDC) estimated that among new cases of childhood diabetes, the proportion of those with type 2 diabetes ranges between eight percent and 43 percent (www.cdc.gov/diabetes/pubs/factsheets/search.htm). Type 2 diabetes begins when the body develops a resistance to insulin and no longer uses insulin properly. Adults and children who develop type 2 diabetes are typically overweight or obese and have a family history of the disease. Furthermore, offsprings of mothers who had diabetes during pregnancy had higher rates of type 2 diabetes and obesity.⁹⁷ It is increasingly recognized that nutrition in early postnatal life may have long-term physiologic effects. Studies have suggested that breastfeeding may be protective against later obesity.^{77,78} Therefore, it seems biologically plausible that there may be a relationship between breastfeeding and long-term glucose and insulin metabolism.

We identified one systematic review by Taylor et al.⁹⁸ in our initial literature search and we also identified one additional systematic review by Owen et al.⁹⁹ during final phase for the preparation of this report. We did not identify any additional systematic reviews.

Commonly considered confounders in the studies of relationship between type 2 diabetes and breastfeeding were age, sex, BMI, birth weight, socioeconomic status, history of parental diabetes, maternal diabetes during pregnancy, maternal diet and smoking, and prepregnancy BMI.

Additional Methodological Comments

We have elected to describe in details only the Owen 2006 systematic review because it superseded the Taylor 2005 systematic review. Owen 2006 included two of the three primary studies that were covered in Taylor 2005. In addition, Taylor 2005 was of poor methodological quality because there was no synthesis of results and it was unclear how conclusions were drawn.

Published Systematic Reviews/Meta-Analyses (Table 16)

Owen 2006 conducted a systematic review and meta-analyses to examine the influence of breastfeeding on type 2 diabetes and blood glucose and insulin concentrations.⁹⁹ Studies that did not provide the odds ratios of type 2 diabetes comparing breastfed to formula-fed subjects were excluded. In addition to two of the three studies included in the Taylor 2005 review (the third study was excluded from Owen 2006 because it did not provide odds ratio of the disease), five additional studies were identified in the Owen 2006 review. Seven studies (six in adults and one in adolescents) totaling 76,744 subjects provided odds ratios that related initial infant feeding methods and type 2 diabetes were included in the meta-analyses. There were three historical cohort, two cross-sectional, one prospective cohort, and one case-control studies. All seven studies were conducted in developed countries. The effects of study size, year of birth, the method of ascertainment of infant feeding status (whether contemporary or recalled up to 71

years after birth), type of formula feeding, study response rate, study design, and whether infants were born pre- or full-term were examined by using meta-regression and sensitivity analyses. Sensitivity analyses were also used to examine the effect of adjustment for important confounders and of fasting status. The methodological quality of this systematic review and meta-analyses was rated grade A.

Six of seven studies related breastfeeding to a lower risk of type 2 diabetes, and there was no evidence of heterogeneity across studies. Overall, the subjects who were breastfed showed a lower risk of type 2 diabetes than those who were formula-fed (pooled adjusted OR 0.61; 95%CI 0.44-0.85, $P=0.003$). Three studies considered the effects of potential confounding by birth weight, parental diabetes, socioeconomic status, and individual or maternal body size, while the other four studies only considered the effects of confounding by age, sex and/or birth weight. However, the odds ratio relating breastfeeding and diabetes risk was similar before (OR 0.55; 95% CI: 0.35-0.86; $P=0.009$) and after adjustment for all the important confounders (OR 0.55, 95% CI 0.34-0.90; $P=0.017$) in the three studies. The method of ascertaining feeding exposure was unrelated to the odds ratios, although there was insufficient power to detect appreciable differences in examining the effect of potential biases (such as study size, year of birth, the method of ascertaining infant feeding status, type of formula feeding, study response rate, study design, and whether infants were born pre- or full-term) by using meta-regression and sensitivity analyses.

Studies Identified after the Published Meta-Analysis/Systematic Review

None was found.

Conclusion

Results from a high-quality systematic review and meta-analyses of seven studies suggest that breastfeeding is associated with a lower risk of type 2 diabetes in later life, compared with formula feeding. Comparing subjects who were ever breastfed to those who were formula fed, the pooled adjusted odds ratio of type 2 diabetes in later life was 0.61 (95%CI 0.44-0.85). However, only three studies appropriately adjusted for all the important confounders, including birth weight, parental diabetes, socioeconomic status, and individual or maternal body size. Even though these three studies found that adjustment did not alter the crude estimate, we cannot be completely confident that potential confounding by birth weight and maternal factors has been ruled out for the overall pooled estimate. This could lead to an overestimate of the association. Publication bias is also a possible explanation for the consistent associations observed in these studies.

Table 16. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and type 2 diabetes

Author Year Population	Study description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitation
Owen 2006 N=76,744 6 studies conducted in adults and 1 study conducted in adolescents	MA of 7 studies (3 historical cohort, 2 cross-sectional, 1 prospective cohort and 1 case-control studies)	Ever breastfed vs. formula fed in all 7 studies. Feeding status was reported as being exclusive in one of these studies.	Age, sex, BMI, birth weight, socioeconomic status, parental diabetes, maternal diabetes in pregnancy, hospital duration in infancy, maternal diet and smoking, and prepregnancy BMI	- Overall, the subjects who were breastfed showed a lower risk of type 2 diabetes than did those who were formula fed (pooled OR: 0.61; 95%CI 0.44-0.85; p=0.003). - OR relating breastfeeding and diabetes risk was similar before (0.55; 95% CI: 0.35-0.86; p=0.009) and after (0.55, 95% CI 0.34-0.90; p=0.017) adjustment in 3 studies that had information on relevant confounders.	A Pooled different study designs
Taylor 2005 N=1,430 People of all ages	Systematic review of 2 cohort studies and 1 case-control study	Variable	Age, sex, birth date, parental diabetes, and birth weight GDM, use of traditional diet, smoking, and alcohol during pregnancy, mother's pregnancy BMI, and birth weight None	- In a retrospective cohort study of 720 people age 10-39 years old, type 2 diabetes (defined by OGTT testing) was 59% less common in exclusively breastfed people compared with those who were exclusively bottle-fed (adjusted OR: 0.41; 95%CI 0.18-0.93) - In a case-control study of native Canadian, Some breastfeeding for one year or longer was a significant independent predictor of future diabetes (adjusted OR: 0.24; 95%CI 0.07-0.84) as was some breastfeeding for 6 months of longer (adjusted OR: 0.36; 95%CI 0.13-0.99). - Type 2 diabetes (defined by OGTT testing) in the next generation was less common among children who were breastfed exclusively for > 2 months (6.9% vs. 30.1% among offspring of women without diabetes and women with diabetes, respectively) than among bottle-fed children (11.9% vs. 43.6%, respectively).	C No synthesis of results and unclear how the conclusions were reached

GDM, gestational diabetes; OGTT, oral glucose tolerance test; BMI, body mass index (kg/m²); SR, systematic review; MA, meta-analysis

Relationship between Childhood Leukemia and Breastfeeding

Background

Leukemia, the most common cancer in children, encompasses multiple diseases including three types: acute lymphocytic leukemia (ALL), acute myelogenous leukemia (AML), and chronic myelogenous leukemia. In the United States, approximately 3,250 children are diagnosed with leukemia each year and 2,400 (74 percent) of them have ALL. With the exception of prenatal exposure to x-rays and specific genetic syndromes, little is known about the causes of childhood ALL (National Cancer Institute: seer.cancer.gov/publications/childhood/leukemia.pdf). Although the majority of human leukemias or lymphomas have no readily identifiable infectious etiologies, viral causes have been identified for Burkitt's lymphoma and a rare form of adult leukemia/lymphoma.^{100,101} Because breast milk is noted for providing passive immunity and protection of newborns from some early infections, investigators have hypothesized that breastfeeding may reduce the risk of childhood leukemia.¹⁰²

Some of the confounders that had been considered in the studies of the relationship between childhood leukemia and breastfeeding were age, sex, race, socioeconomic status (such as mother's education, region of residence at the time of diagnosis, parental occupation, deprivation index, and/or annual household income), parental alcohol consumption, parental smoking, birth order, birth weight, parity, and age of mother at birth of index child.

Additional Methodological Comments

We identified four systematic reviews or meta-analyses that examined the relationship between breastfeeding and childhood leukemia¹⁰³⁻¹⁰⁶. We have elected to describe in details only the Guise 2005¹⁰⁴ systematic review and Kwan 2004¹⁰⁶ meta-analysis because they superseded the Beral 2001¹⁰³ meta-analysis and Davis 1998¹⁰⁵ systematic review. All studies included in Beral 2001 and Davis 1998 were also included in Kwan 2004 or Guise 2005, except that Guise 2005 review excluded studies published before 1990, unpublished data, and studies conducted in developing countries. Kwan 2004 also included three additional studies published after 2001. In addition, the reporting and analysis in Beral 2001 and Davis 1998 were poor due to deficiency in reporting of the meta-analysis methods, search strategy and/or study eligibility criteria, and there was a lack of consideration of potential confounding in the included studies.

Guise 2005's systematic review was the only study that reported methodological grading for individual studies. Guise 2005 did not perform a meta-analysis. Kwan 2004 combined all case-control studies regardless of their study quality into a meta-analysis. To understand further how study quality can affect the effect estimate, we decided to perform meta-analyses on the effect estimates from the four studies graded as good or fair quality in Guise 2005's systematic review.

Published Systematic Reviews/Meta-Analyses (Table 18)

Guise 2005 conducted a systematic review of case-control studies related to breastfeeding and the risk of childhood leukemia. Ten case-control studies totaling 9,653 subjects with leukemia were included. All the studies were conducted in developed countries. Study quality

was rated by the authors in a three-levels scale: good, fair, or poor. The following aspects of the study quality were assessed: reliability of the diagnoses of leukemia, comparability of the case and control groups, differences in nonrespondents in cases versus controls, and the conduct of controlling for confounding. There were two good, two fair, and six poor quality studies. All studies but one focused solely on childhood leukemia. There was no meta-analysis or statistical analysis performed. The methodological quality of this systematic review was rated grade A.

Many of the studies in Guise 2005 systematic review did not provide data on exclusivity of breastfeeding and did not consider potential confounders such as infectious exposures from household or school contacts. Six of the ten studies explicitly sought to characterize the relationship between breastfeeding and leukemia.

Guise 2005 concluded that the few high-quality studies disagreed in regards to the association between breastfeeding and the risk of ALL. Specifically, the two good-quality studies, or UKCCS and CCG study,^{103,107} presented “conflicting results” (Table 17). Similarly, the two fair-quality studies disagreed on the protective effect of breastfeeding.

Kwan 2004 conducted a meta-analysis of 14 case-control studies on breastfeeding and the risk of childhood leukemia. A total of 8,051 subjects with leukemia were included in the analysis. The meta-analysis examined the relationship between short-term breastfeeding (defined as breastfeeding for 6 months or less) or long-term breastfeeding (defined as breastfeeding for more than 6 months) and the risk of childhood leukemias. Twelve studies from developed countries and two studies from developing countries, including any type of leukemia in children 15 years or younger that reported odds ratio and duration of breastfeeding, were included in the meta-analysis. Kwan 2004 did not formally assess the individual study quality. However, the potential for confounding in each study was considered in the meta-analysis. The methodological quality of this meta-analysis was rated grade A.

For each of the 14 primary studies in the meta-analysis, Kwan 2004 selected odds ratio adjusted for SES when available. The meta-analysis showed a statistically significant reduced risk of ALL with short- and long-term breastfeeding (OR 0.88, 95%CI 0.80 - 0.96; OR 0.76, 95%CI 0.68 - 0.84, respectively). The analysis reported a statistically significant reduction in AML for long-term breastfeeding (OR 0.85, 95%CI 0.73 - 0.98) but not for short-term breastfeeding (OR 0.90, 95%CI 0.80 - 1.02). The unadjusted and adjusted odds ratios for reduction of the risk of ALL in short-term breastfeeding were both statistically significant. For reduction of the risk of ALL in long-term breastfeeding, only the adjusted odds ratio was statistically significant. The same sensitivity analyses were performed for the reduction of the risk of AML in short- and long-term breastfeeding, and similar results were found.

Studies included in the additional analysis. Guise 2005 identified two good and two fair methodological quality studies. We have analyzed these studies further as detailed below.

The two good studies were the UKCCS and the CCG studies. The UKCCS included 1,401 (87 percent) ALL and 214 (13 percent) AML case patients recruited from health programs that enrolled 98 percent of the total childhood cancers throughout England, Scotland and Wales over the period 1991-1998; control subjects were selected from population-based health rosters. In contrast, the CCG study included 1,744 (79 percent) ALL and 456 (21 percent) AML case patients enrolled from specific CCG centers, and control subjects were selected via random-digit-dialing. Both studies excluded leukemia diagnoses less than 1 year of age because most leukemias occurring during infancy are thought to have different etiologies from childhood leukemias. Both studies found that long-term breastfeeding (> 6 months duration) was protective for ALL, but the confidence interval for the risk estimate from the UKCCS did not exclude unity

(1.00), whereas the confidence interval for the risk estimate reported by the CCG study clearly excluded unity (1.00), indicating statistical significance. The UKCCS, however, found no association between short- or long-term breastfeeding and the risk of AML, while the CCG study found a significant protective effect of long-term breastfeeding for AML.

The two fair quality studies were Dockerty 1999 and Rosenbaum 2000. They were graded fair quality due to potential selection biases.^{108,109} Dockerty 1999 included 121 newly diagnosed leukemia cases (ages 0-14 years) and 121 age- and sex-matched control subjects selected randomly from the New Zealand national birth records. The primary purpose of the study was to examine the relationship between infections, vaccinations, and the risk of childhood leukemia. Breastfeeding was one of the secondary factors examined in the study. Compared with children who never breastfed, those who breastfed for more than 6 months to 1 year had about a 20 percent reduced risk of ALL; those who were breastfed for more than 1 year had the lowest risk (OR 0.47; 95%CI 0.15 – 1.43). Even though these estimates were not statistically significant (Table 22), a trend analysis indicated a statistically significant effect in reducing the risk of ALL with increasing duration of breastfeeding (P = 0.04). Rosenbaum 2000 included 255 ALL cases from hospital registries and 760 matched control randomly selected from birth certificates in US. This study aimed to examine the relationship between early child-care (including breastfeeding) and the risk of childhood ALL. Like Dockerty 1999, children under 1 year of age were included in Rosenbaum 2000. The analysis of the relationship between breastfeeding and the risk of childhood ALL did not adjust for any other potential confounders, except for the factors used to identify matched controls (gender, race, and birth year). They found that 47 percent of cases of ALL and 51 percent of control were breastfed at birth (OR 1.20, 95%CI was not reported). This association was not statistically significant.

Meta-analysis. We used a random-effects model to combine SES-adjusted odds ratios of ALL in relation to short-term (≤ 6 months) and long-term (> 6 months) breastfeeding from UKCCS¹⁰³, CCG¹⁰⁷ study, and Dockerty 1999 (Table 17). Rosenbaum 2000 was excluded from the analysis because the duration of breastfeeding was not reported. The results from our meta-analysis suggest that long-term breastfeeding is associated with a reduction in the risk of ALL (OR 0.80; 95%CI 0.71 - 0.91).

Table 17. Combined SES-adjusted ORs of ALL for the three case-control studies rated as good and fair methodological quality in the systematic review by Guise et al. (2005)

Study	OR (≤ 6 mo vs. never BF)	Lower CI	Upper CI
UKCCS 2001			
<1 mo	0.90	0.77	1.04
1-6 mo	0.98	0.82	1.17
CCG study 1999	0.86	0.73	1.01
Dockerty 1999	1.24	0.47	3.23
Pooled	0.91	0.83	1.00
Kwan 2004 meta-analysis results	0.88	0.80	0.96

Study	OR (> 6 mo vs. never BF)	Lower CI	Upper CI
UKCCS 2001	0.89	0.75	1.05
CCG study 1999	0.72	0.60	0.87
Dockerty 1999			
>6 mo to 1 yr	0.82	0.29	2.27
>1 yr	0.47	0.15	1.43
Pooled	0.80	0.71	0.91
Kwan 2004 meta-analysis results	0.76	0.68	0.84

Conclusion

Our meta-analyses of the three case-control studies concerning breastfeeding and the risk of ALL were consistent with the results from Kwan 2004's meta-analysis, but with smaller effect size and smaller statistical significance (Table 17). Kwan 2004 also found an association between a history of breastfeeding and a reduced risk of AML. We conclude that there is association between a history of breastfeeding of at least 6 months duration and a reduction in the risk of both ALL and AML.

Further evaluation of the biological mechanisms underlying this relationship while taking into consideration potential biases can be achieved with more large-scale case-control studies utilizing population-based and socioeconomic status-matched controls.

Table 18. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and childhood leukemia

Author Year Population	Study description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitations
Guisse 2005 Childhood ALL or all childhood leukemias: 9,653	SR of 10 case- control studies in developed countries	Any measures of breastfeeding and comparators	Gender, year of birth, race, SES (region of residence, maternal education, family income), gestational age, birthweight, maternal age, and smoking	- Six studies were conducted in European countries. Six of the studies explicitly sought to characterize the relationship between breastfeeding and leukemia as the primary objective, whereas the others included breastfeeding measures from the perspective measuring broader characteristics of the immune system and early infections in the etiology. Of the 10 studies, 2 were of good quality, 2 were of fair quality, and 6 were of poor quality. - The 2 good-quality studies (UKCCS and CCG studies) present “conflicting results” regarding the association between breastfeeding and leukemia. - Similarly, the 2 fair-quality studies disagreed on the protective effect of breastfeeding.	A Interpretation of “conflicting results” for the 2 good-quality studies was based on statistical significance only
Kwan 2004 Children with ALL, AML or ANLL: 6,835 (Note: 8 of the 14 studies excluded cases of leukemia in infants)	MA of 14 case- control studies in both developed and developing countries	Compared short-term breastfeeding (≤ 6 months) or long-term breastfeeding (> 6 months) to no breastfeeding or never been breastfed	Age, sex, race, SES (mother’s education, region of residence at the time of diagnosis, parental occupation, deprivation index, and/or annual household income), parental alcohol consumption, parental smoking, birth order, birthweight, parity, age of mother at birth of index child.	- A significant negative association was observed between short-term breastfeeding and ALL (OR=0.88, 95% CI 0.80-0.96), but the AML results (OR=0.90, 95% CI 0.80-1.02) were not significant. - A significant negative association was observed between long-term breastfeeding and ALL (OR=0.76, 95% CI 0.68-0.84), and AML (OR=0.85, 95% CI 0.73-0.98).	A 2 studies from developing countries
Beral 2001 Children with all leukemia, including ALL: 7,401	MA of 15 case- control studies in both developed and developing countries	Compared ever breastfeeding, breastfeeding duration ≤ 6 months, or breastfeeding duration > 6 months to never been breastfed	ND	There is evidence of a statistically significant reduction in the OR associated with ever having been breastfed (OR=0.86, 95%CI 0.81-0.92) and having been breastfed for more than 6 months (OR=0.78, 95% CI 0.71-0.85).	C No methods of meta- analysis; no description of search strategy and study eligibility; no consideration of potential confounding

Table 18. Continued

Author Year Population	Study description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitations
Davis 1998 Children with all leukemia, ALL or ANLL: 1,656	SR of 5 case- control studies in developed countries	Compared any breastfeeding, or long- term breastfeeding (6-8 months) to artificial infant feeding (or never been breastfed)	None	None of the included studies reported a significant association between all leukemia, ALL or ANLL and infant feeding.	C Poor reporting of search strategy and study eligibility; no consideration of potential confounding
ALL, acute lymphocytic leukemia; ANLL, Acute non-lymphocytic leukemia; AML, acute myeloid leukemia; SES, socioeconomic status; SR, systematic review; MA, meta-analysis					

Relationship between Infant Mortality and Breastfeeding

Background

Infant mortality (both neonatal and post-neonatal mortality) is on the decline in both developing and developed countries during the past four decades. In the United States, the national average of infant mortality was 7.0 deaths per 1000 live births in 2002.¹¹⁰ Leading causes of infant death in a developed country include congenital abnormalities, pre-term births, low birth weight, Sudden Infant Death Syndrome (SIDS), problems related to complications of pregnancy, and respiratory distress syndrome.¹¹⁰ The common risk factors that are associated with infant mortality include infant and maternal age, gender, race, socioeconomic status, birth order, birth weight, congenital malformation at birth, and maternal smoking during pregnancy. Breastfeeding protects against infectious diarrhea and respiratory diseases, which are the leading causes of infant mortality in developing countries.¹¹¹ However, the role of breastfeeding in preventing infant deaths in developed countries is less clear. We elect to review the relationship between breastfeeding and post-neonatal mortality.

Additional Methodological Comments

For this section, infant mortality is defined as any death that occurred more than 1 month but less than 12 months after birth. Neonatal mortality (death < 1 month of age) is not considered in this review. No published meta-analysis was identified that evaluated the relationship between breastfeeding and infant mortality. We included only studies conducted in developed countries that evaluated the relationship between breastfeeding and infant mortality (excluding SIDS).

Results (Table 19)

We identified two studies conducted in the United States that evaluated the role of breastfeeding in infant deaths other than SIDS.^{112,113}; one study did not qualify for inclusion in this review because it was an analysis based on data collected from cross-sectional surveys.¹¹³ The study that qualified for inclusion was a case-control study that selected subjects from nationally representative samples.¹¹² The study excluded infants who died at less than one month of age. The methodological quality of the study was rated grade B.

Chen 2004 analyzed the 1988 US National Maternal and Infant Health Survey (NMIHS) data using a case-control study design. Their sample included 1,204 infants who died between 28 days and 1 year of age. The authors excluded infants who died from congenital anomalies or malignant tumors. In an attempt to address the issue of reverse causality, only infants who survived beyond the neonatal period were included and feeding status was categorized based on the assessments undertaken some time before death occurred. The controls included 7,740 live births who were alive and older than 1 year of age at the time of survey. Mothers were surveyed using a mailed questionnaire. Breastfeeding was assessed as “ever breastfed” or “never breastfed.” The study did not report the mean or the range of duration of maternal recall for breastfeeding. Overall and cause-specific odds ratios were calculated to evaluate “ever” versus “never” breastfeeding. Since there was an oversampling of black and low birth weight infants in the original sample, in order for the study population to be more representative of the US

population, additional analyses were performed with SUDAAN software adjusted samples. This adjusted sample included a total of 9,145 cases and 3,186,497 controls. The analyses were adjusted for sampling strategy and potential confounders such as maternal age, education, smoking, infant gender, race, birth weight, congenital malformation, live birth order, plurality, and status of enrollment in the Special Supplemental Nutrition Program for Women, Infant, and Children program (WIC). The study reported that the odds of death in the postneonatal period were 21 percent lower for “ever breastfed” compared with “never breastfed” infants. However, in subgroup analyses of cause-specific death, the only statistically significant association was reported between SIDS (in the original sample) or injury-related death (in the SUDAAN-adjusted sample) and “never breastfed” status.

Conclusion

One study of moderate methodological quality using a large sample of infants reported a protective effect of breastfeeding in reducing infant mortality after controlling for some of the potential confounders. However, in subgroup analyses of the study, the only statistically significant association reported was between “never breastfeeding” and SIDS or the risk of injury-related deaths. Because of the limited data in this area, the relationship between breastfeeding and post-neonatal infant mortality remains unclear. Further research is warranted.

Table 19. Summary of case-control study on the relationship between breastfeeding and infant mortality

Author year <i>Country</i>	Cases (N)	Control (N)	Definition of infant mortality	Mean Age at Dx (year)	Breastfeeding group	Comparator group	OR* (95% CI)		Potential confounders adjusted	Quality and limitations
							Crude	Adjust		
Chen 2004 USA National Maternal and Infant Health Survey 1998 (NMIHS)	Postneonatal death ^a (9145)	Live birth ^a (3 186 497)	Death	> 28d of age	% Ever	% Never	nd	0.79 (0.67-0.93)	Maternal: age, education, smoking Infant: gender, race, birth weight, congenital malformation, live birth order, plurality, and WIC status	B (in subgroup analyses statistically significant association was reported between injury-related death or SIDS and "never breastfed" status)
	Postneonatal death ^b (1204)	Live birth ^b (7740)					nd	0.76 (0.65-0.88)		

^a SUDAAN software adjusted sample

^b Original sample

Relationship between Sudden Infant Death Syndrome (SIDS) and Breastfeeding

Background

Sudden infant death syndrome (SIDS) is the leading cause of mortality among infants aged 1 to 12 months in the United States.¹¹⁴ SIDS accounted for a death rate of 0.55 per 1000 live births for the year 2004 according to the National Center for Health Statistics (NCHS) at CDC (www.cdc.gov/nchs/deaths.htm). Several modifiable risk factors for SIDS are sleeping positions, maternal smoking, and bed sharing. Other potential risk factors include birth weight, gender, and socioeconomic status. The relationship of breastfeeding and SIDS has been evaluated among a broad range of potential risk factors. However, the role of breastfeeding as a protective factor in SIDS is unclear. One meta-analysis published in 2000 assessed the relationship of breastfeeding and SIDS.

Additional Methodological Comments

We identified four eligible studies conducted in developed countries that evaluated the role of breastfeeding in SIDS, and published since 1997, the cutoff date for the literature search of the published meta-analysis. .

Published Systematic Review/Meta-Analysis (Table 20)

One meta-analysis¹¹⁵ of 23 studies (18 case-control; four nested case-control; and one observational cohort) evaluated the relationship of breastfeeding and SIDS. All studies were conducted in developed countries from 1965 to 1997. Studies that reported a minimal definition of SIDS – sudden unexplained death of an infant or young child – met the eligibility criteria. The meta-analysis analyzed a total of 4,251 cases of SIDS and 58,055 controls. Seventeen studies included subjects from birth to 2 years old when SIDS occurred; eight studies did not provide age data. The studies differed in their definition of breastfeeding exposure. Also, the studies varied in their description of SIDS. Of these, only 14 studies reported autopsy-confirmed diagnoses of SIDS. Three studies were specifically designed to examine the relationship of breastfeeding and SIDS. The rest of the studies examined multiple risk factors and their association with SIDS, of which feeding history would be one.

The meta-analysis utilized a random effects model. It reported an overall risk of SIDS twice as great for formula-fed infants compared with breastfed infants (crude odds ratio (OR) of 2.11; 95% CI 1.66 to 2.68). The methodological quality of individual studies was appraised. The authors conducted a separate meta-analysis for those studies published since 1988 when more advanced epidemiological and autopsy procedures were available. A separate analysis was also performed for those studies with “high” quality scores. The results from these meta-analyses concurred with the overall result. Heterogeneity was not explored, i.e., no subgroup analyses were performed to account for the different definitions of interventions or outcomes. The authors reported “no publication bias”. Differences in case matching precluded them from combining adjusted odds ratios in the meta-analysis. Four of the 16 studies showed a dose response trend

with the risk of SIDS increasing with increasing formula feeding. The overall methodological quality of the meta-analysis was rated grade C.

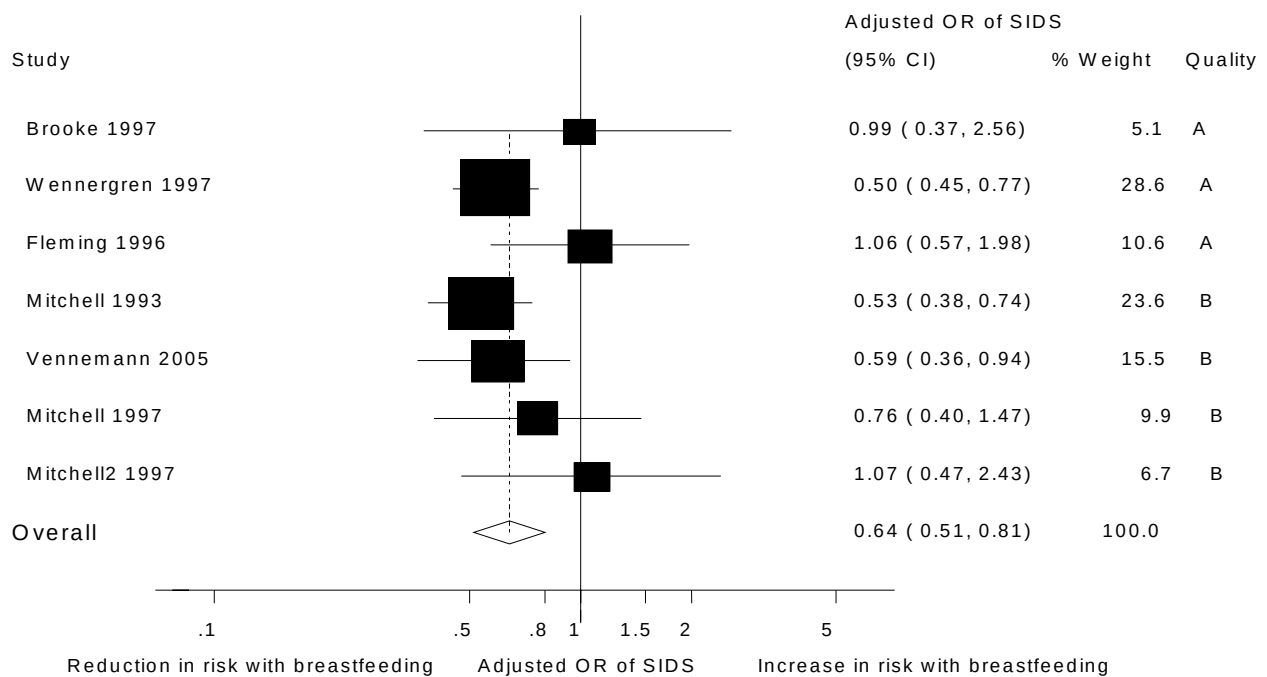
Studies Identified after the Published Meta-Analysis (Table 21)

We identified four eligible studies from five publications conducted in developed countries and published since 1997 that evaluated the relationship of breastfeeding and SIDS.¹¹⁶⁻¹²⁰ Three were case-control studies and the fourth, a case-cohort study (a cohort study analyzed as a case-control study). The methodological quality ranged from grades B to C. There were a total of 769 cases of SIDS and 2,681 controls.

All studies were designed to evaluate a broad range of potential risk factors for SIDS. The studies differed in their description of the duration of breastfeeding. Similarly, the studies varied in their definition of the time interval when SIDS occurs in infants, but all studies reported autopsy-confirmed diagnoses. The mean age of the infants with SIDS ranged from 2 to 19 weeks. All studies provided adjusted odds ratios for the association of breastfeeding and SIDS. Three of the four studies identified statistically significant increased risk of SIDS in bottle-fed infants.^{117,119,120} Two studies reported that the risk of SIDS was twice or more for non-breastfed infants compared with some or ever breastfed infants.^{117,119} One study reported an approximately two times increased risk of SIDS among those breastfed less than 2 weeks compared with those breastfed more than 2 weeks.¹²⁰ One case-cohort study (a cohort study analyzed as case-control study) did not find a statistically significant increased risk of SIDS in bottle fed infants.¹¹⁸

Meta-analysis results. Because of the limitations of the previous meta-analysis, we elected to conduct our own meta-analysis using only studies that provided an objective definition of SIDS (autopsy confirmed SIDS among infants 1 week to 1 year of age), clear reporting of breastfeeding data, and outcomes adjusted for important confounders or risk factors (e.g., sleeping positions, maternal smoking, and socioeconomic status). Four studies included in the previously published meta-analysis¹²¹⁻¹²⁴ and two studies published since 1997 met the eligibility criteria.^{118,120} The majority of the studies provided data on ever versus never breastfeeding and this was combined using a random effects model. The results from our meta-analysis found that ever breastfeeding was associated with a reduction in both crude and adjusted risk of SIDS (crude OR 0.41; 95%CI (0.28, 0.58), and adjusted OR 0.64; 95%CI (0.51, 0.81), respectively); both estimates were statistically significant with a reduction in SIDS for the ever breastfed infants.

Figure 10. Random effects model of summary estimate evaluating the association of breastfeeding and SIDS



Conclusion

Results from the previously published meta-analysis of case-control studies concluded that an overall crude risk of SIDS was twice as great for formula-fed infants compared with breastfed infants. The conclusion may be biased because the reported association was not adjusted for potential confounders. Misclassification biases may occur because of differences among studies with regard to definitions of breastfeeding exposure, definitions of SIDS, and the wide age range of population included in the studies.

Findings from the four studies published subsequent to the meta-analysis in developed countries concurred with the findings from the meta-analysis. All studies reported autopsy-confirmed diagnoses of SIDS and adjusted for potential confounders. However, the definitions of breastfeeding exposure and the time intervals accepted for defining SIDS varied across studies. Three of four studies reported statistically significant increased risk of SIDS associated with non-breastfeeding or reduced duration of breastfeeding and the fourth study reported a statistically non-significant increased risk.

Our meta-analysis included only studies that reported clear definitions of exposure, outcomes, and results adjusted for well-known confounders or risk factors for SIDS. The summary estimate found a statistically significant adjusted odds ratio for an association between breastfeeding and a reduced risk of SIDS (adjusted OR 0.64, 95%CI 0.51 - 0.81). We conclude that there is a relationship between breastfeeding and a reduced risk of SIDS.

Table 20. Summary of systematic review/meta-analysis on the relationship between breastfeeding and Sudden Infant Death Syndrome (SIDS)

Author year	Study description	Number of subjects	Population	Intervention /Comparator	Results	Quality of SR/MA
McVea 2000	MA of 23 observational studies	Cases 4,251 Controls 58,055	Children who were diagnosed of SIDS	Breastfeeding (any)	The pooled OR for the 23 studies using random effects model resulted in an OR = 2.11 (95% CI 1.66-2.68) i.e., the overall risk of SIDS was twice as great for bottle-fed infants compared to breastfed infants. The pooled OR from the higher quality studies also demonstrated a two fold increase in risk among bottle fed infants OR = 2.24 The pooled OR from studies after 1988 OR = 2.32 Confounders: Individual studies adjusted for potential confounders; 6 studies reported adjusted OR; 4 studies reported no protective effect of breastfeeding, while 2 reported adjusted OR that remained significant. Dose-response relationship: 4 out of 9 studies showed a dose response trend with the risk of SIDS increasing with increasing formula feeding. None of the studies had sufficient power to demonstrate a statistically significant difference between partial vs. no breastfeeding.	C

SR, systematic review; MA, meta-analysis

Table 21. Summary of case-control studies on the relationship between breastfeeding and sudden infant death syndrome (SIDS) included in the meta-analysis

Author year Country	Cases (N)	Control (N)	Definition of SIDS	Mean Age at Dx (week)	Breastfeeding group	Comparator group	OR* (95% CI)		Potential confounders adjusted	Quality and limitations
							Crude	Adjust		
Wennegran, 1997 Alm 2002 Scandinavia	244	869	Validated definition	16	Ever (At the time of death)	Never	0.59 (0.4, 0.83)	0.5 (0.45, 0.77)	Sleep position, maternal smoking, bottle feeding at the time of death and age at the time of death	A
Fleming, 1996 UK	195	780	7d-1y	nd	Ever	Never	0.5 (0.35, 0.71)	1.06 (0.57, 1.98)	Maternal age, gestation, birth weight, exposure to tobacco, sleeping environment	A
Brooke, 1992 UK (Scotland)	147	275	7d-1y	15-18	Ever (At the time of death)	Never	0.22 (0.11, 0.47)	0.99 (0.37, 2.56)	Exposure to Parental smoking, sleep position, old mattress use, maternal age <27, deprivation score of 7, drug treatment in previous wk, marital status of the mother, SE status, gender of the infant, birth weight etc	A
Mitchell, 1993 New Zealand Population 1987-1990	460	1757	>28d-1yr	nd	Ever (At the time of discharge from obstetric dept)	Never	0.42(0.33, 0.53)	0.53(0.38, 0.74)	Region, time of day, baby's age, antenatal class, school leaving age of mother, marital status of mother, sex of baby, admission to neonatal unit, number of previous pregnancies, socioeconomic status, birthweight, gestational age, race of baby, season, mothers age at first pregnancy, mothers age at birth, sleeping position, bed sharing with another person, maternal smoking and breastfeeding	B
Vennemann 2005 Germany	333	998	8d-1 yr	19	>2 wk	<2 wk	0.19 (0.14, 0.25)	0.59 (0.36, 0.94)	Maternal age, family status, smoking in pregnancy. Previous live births and socioeconomic status	B
Mitchell 1997 ^a New Zealand	79	679	29d-1 yr	2.6-9	Any at initial contact	None at initial contact	0.60 (0.35, 1.03)	0.76 (0.40,1.47)	Maternal age, marital status, age mother left school, previous number of pregnancies, infant's sex, ethnicity of infant, birthweight, sleep position, and bed sharing/maternal smoking combinations	B
	38	588			Any at 2 mo	None at 2 mo	0.76 (0.41, 1.39)	1.07 (0.47, 2.43)		

^a A case-cohort study

Part II. Preterm Infant Outcomes

Relationship between Necrotizing Enterocolitis and Breast Milk Feeding in Preterm Infants

Background

Necrotizing enterocolitis (NEC) is a serious gastrointestinal disease in the preterm infants. No definitive causes have been identified. A population-based epidemiological study published in 2002 reported that the highest incidence occurred in infants with birth weights 750 to 1000 g and decreased with increasing birth weights.¹²⁵ Observational studies have suggested that breast milk might be protective. There have been very few randomized controlled trials (RCTs) that examined this issue. McGuire 2001 performed a meta-analysis of RCTs of breast milk comparing with formula milk in preterm infants to reduce the risk of NEC.¹²⁶ Since that meta-analysis, we identified one RCT¹²⁷ and two prospective cohort studies^{128,129} that reported outcome of NEC in preterm infants in relation to a history of breast milk feeding. Some of the potential confounders that may affect the results of neonatal morbidity and mortality include birth weight, ethnicity, and sex.¹²⁸

Published Systematic Review/Meta-Analysis (Table 23)

McGuire 2001's meta-analysis compared formula feeding with term breast milk feeding in low birth weight or preterm infants. Three RCTs published in 1983 and 1984 totaling 308 preterm infants were included: Gross 1983 compared formula with unfortified term donor breast milk;¹² Tyson 1983 compared preterm formula with pooled banked term breast milk;¹³ and Lucas 1984¹³⁰ (results for NEC reported in 1990¹³¹) compared preterm formula with banked term breast milk as the sole diet. In the meta-analysis comparing formula with breast milk, the risk ratio for developing NEC was 2.5 (95% CI 0.9 - 7.3); risk difference was 0.05 (95% CI 0.00 - 0.1). The authors concluded that there was no statistically significant difference in the risk of NEC with either form of milk feeding.

Studies Identified after the Published Systematic Review/Meta-Analysis (Table 24)

Randomized controlled trial. Schanler 2005 enrolled 243 infants $\leq$ 29-week gestation whose mothers were expected to breastfeed.¹²⁷ If these infants' own mothers' milk were unavailable, the infants were then randomly assigned to receive either pasteurized donor milk or preterm formula. However, both groups continued to receive mother's milk partially if they were available during the study. The infants who were fed mother's milk exclusively were not randomized and served as a reference group. The incidence of NEC in Donor milk versus Preterm formula was 5/78 versus 10/88 ($P = 0.27$). The non-randomized group "mother's milk" had fewer repeated episodes of late-onset sepsis and/or NEC (OR 0.18, 95% CI 0.04 - 0.79) compared with combined groups "donor milk" and "preterm formula". The methodological quality of this study was rated B.

Prospective cohort. Furman 2003 was a prospective cohort study on 119 infants with gestational age < 33 weeks and birth weight 600-1499 g.¹²⁸ Enteral feeding was begun by day 2 or 3 of life, parenteral nutrition was continued until a daily enteral intake of 120 mL/kg of body weight was reached. Infants received their mother's milk in the sequence it was expressed, except that fresh rather than frozen milk was given if available. Maternal milk was fortified, and preterm infant formula was offered when the infant reached a daily oral intake of at least 110 mL/kg. Limited availability of maternal milk was the sole reason infants were fed preterm formula in addition to maternal milk. Four subgroups were analyzed: no maternal milk, daily maternal milk of 1-24 mL/kg, 25-49 mL/kg, and ≥ 50 mL/kg. Rates of NEC did not differ according to the amounts of maternal milk received. The results of the regression analysis were adjusted for birth weight, ethnicity, and sex. The methodological quality of this study was rated B.

Rønnestad 2005 was a prospective cohort study of late-onset sepsis on 462 infants with gestational age <28 weeks or birth weight < 1000 g in Norway.¹²⁹ NEC was not the primary outcome of interest; it was studied as a potential confounder in the analysis of late-onset sepsis. Four hundred five survived until day 7. Participating centers had a common policy of achieving full enteral feeding with the mother's milk or banked donor milk as early as possible, although there was no uniformity in a detailed protocol for feeding strategies. Enteral feeding with breast milk was commenced within 1, 2, or 3 days for 61 percent, 92 percent, and 96 percent of the infants, respectively. Nine of 405 (2.2 percent) patients had confirmed NEC. There was no concurrent comparison reported in this study. The methodological quality of this study was rate C with respect to the outcome of NEC.

Updating the previous meta-analysis. We performed a new meta-analysis using a random effects model by combining the data from the Schanler 2005 RCT with the three RCTs in McGuire 2001. We combined all breast milk into one group because the proportion of the preterm versus term banked breast milk in the four studies cannot be determined. For outcome, we only counted confirmed cases of NEC as provided by the authors (either pneumatosis intestinalis or confirmed at surgery). We reported the results as risk ratios of developing NEC. The meta-analysis of four RCTs with a total of 476 infants provided a risk ratio of 0.42 (95% CI 0.18, 0.96) for the development of NEC, in favor of breast milk (Table 22; Figure 11).

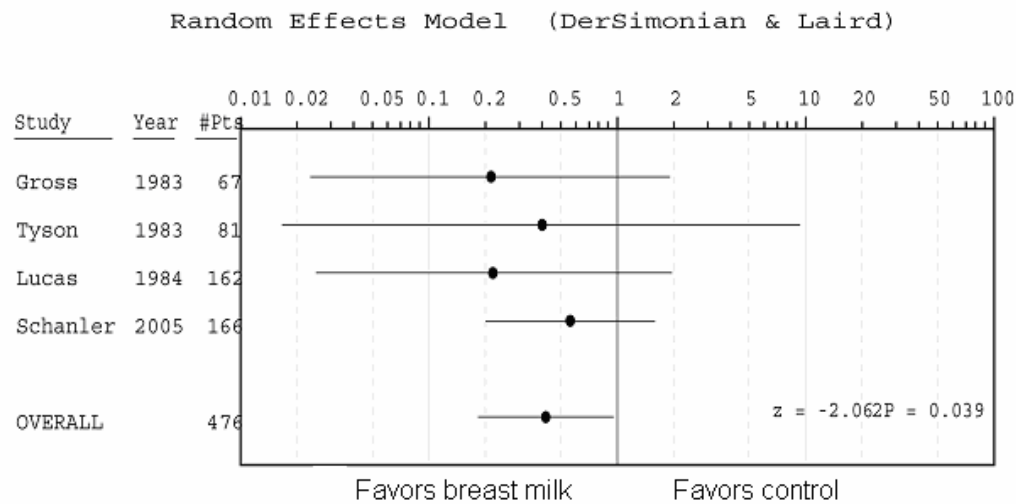
Table 22. Meta-Analysis of four RCTs on the effects of breast milk feeding and NEC in preterm infants: Random Effects Model (D&L)

Study, year	Breast milk feeding		Control		Risk Ratio	95% CI	
	Event	Total	Event	Total		Low	High
Gross, 1983	1	41	3	26	0.21	0.02	1.93
Tyson, 1983	0	37	1	44	0.39	0.02	9.41
Lucas, 1984	1	86	4	76	0.22	0.03	1.93
Schanler, 2005	5	78	10	88	0.56	0.20	1.58
Total patients = 476	7	242	18	234	0.42	0.18	0.96

$z = -2.0629$ $2P = 0.039$

Overall Heterogeneity: $Q = 1.02$ $Tau^2 = 0.0000$

Figure 11. Meta-Analysis of four RCTs on the effects of breast milk feeding and NEC in preterm infants



Conclusion

Even though the observational study by Furman et al. did not find a difference in the rates of NEC according to the amount of maternal milk received by the infants, our meta-analysis of four RCTs demonstrated that there was a marginally statistically significant association between breast milk feeding and the reduction in the risk of NEC. The confidence interval for the estimate in the relative risk reduction ranged from four percent to 82 percent. The absolute risk difference was five percent. The wide confidence of the estimate reflects the relatively small number of total subjects in the studies and the small number of events. One must be cognizant of the clinical heterogeneity underlying these RCTs in interpreting the findings of the meta-analysis. Some of them were: different time periods when the studies were conducted; different preterm formulas as comparators; wide range of gestational ages and birth weights in the subjects; different degree of illnesses in the subjects; and others. How the heterogeneity in the studies affected the findings is unclear. Lastly, one may question the importance of an absolute risk difference of five percent between groups. Taking into account the high case-fatality rate of NEC, we consider this estimate is of meaningful clinical difference. In conclusion, there is evidence to support an association between breast milk feeding and a reduction in the risk of NEC in preterm infants.

Table 23. Summary of systematic review/meta-analysis on the relationship between breast milk feeding and necrotizing enterocolitis (NEC) in preterm infants

Author year	Study description	N	Population	Intervention/Comparator	Results	Quality for SR or MA
McGuire 2001	MA of 3 RCTs	310	Gross 1983 27-33 wk, < 1600 g N=67 US	Unfortified term donor breast milk, fed until 1800 g or until withdrawal secondary to feed intolerance or NEC; compared to standard calorie, protein enriched formula	formula vs. breast milk, RR: 2.5 (95% CI 0.9, 7.3); RD: 0.05 (95% CI 0.0, 0.1)	B
			Tyson 1983 "very low birth weight infants" N=81 US	Pooled banked term breast milk, allocation at 10 th day of life, fed until 2000 g or until withdrawal secondary to illness requiring parenteral fat or protein; compared to enriched calorie and protein formula		
			Lucas 1984 preterm infants < 1850 g N=162 UK	Banked term breast milk 200 mL/kg/d, fed until 2000 g or until d/c; compared to enriched calorie and protein formula		

SR, systematic review; MA, meta-analysis

Table 24. Summary of studies on the relationship between breast milk feeding and necrotizing enterocolitis (NEC) in preterm infants

Author year	N	Population	Intervention/Comparator	Outcomes	Quality
Randomized Controlled Trial					
Schanler 2005	243	<29 wk infants, mothers expected to breastfeed US	Donor milk (DM) (+ mother's milk partially) Small quantities of mother's milk (~20 mL/kg/day) was initiated in the first week and continued for ~3 to 5 days before the volume was advanced. Milk intake was increased by ~20 mL/kg daily to 100 mL/kg, at which time milk fortifier was added, this was advanced by ~20 mL/kg daily until 160 mL/kg per day was achieved. Comparator: Preterm formula (PF) (+ mother's milk partially) Reference group: non-randomized, exclusive mother's milk (MM)	Incidence of NEC: DM vs. PF: 5/78 vs. 10/88 (P=0.27) Non-randomized group MM had fewer repeated episodes of late-onset-sepsis and/or NEC (OR 0.18; 95% CI 0.04–0.79) compared with combined groups DM and PF	B
Prospective Cohort Studies					
Furman 2003	119	<33 wk, birth weight 600-1499 g US	The study compared the effect of varying dosages of maternal milk on neonatal outcomes. Intravenous dextrose during the first 24 hrs; enteral intake was begun by day 2 or 3 of life, parenteral nutrition was continued until a daily enteral intake of 120 mL/kg of body weight was reached. Infants received their mother's milk in the sequence it was expressed, except that fresh rather than frozen milk was given if available. Maternal milk was fortified, and preterm infant formula was offered when the infant reached a daily oral intake of at least 110 mL/kg. Limited availability of maternal milk was the sole reason infants were fed preterm formula in addition to maternal milk.	0 maternal milk 3/40 1-24 mL/kg 2/29 (OR 1.15 95% CI 0.8-12.13) 25-49 mL/kg 2/18 (OR 1.99 95% CI 0.14-21.03) ≥ 50 mL/kg 0/32 (OR 0 95% CI 0 – 3.56) Results were adjusted for birth weight, ethnicity, and sex	B
Rønnestad 2005	462	<28 wk or birth weight < 1000g Norway	Tube feeding with breast milk was usually started within a few hours after delivery, with 1 to 2 mL of milk every 2 or 3 hrs, increasing by 0.5 to 1 mL every 6 to 8 hrs as tolerated. Enteral nutrition was supplemented with parenteral glucose from day 1, amino acids and lipids from day 2 and day 3, respectively.	Enteral feeding with breast milk was commenced within 1, 2, or 3 days for 61%, 92%, and 96% of the infants. 9/405 (2.2%) patients had confirmed NEC.	C No adjustment for potential confounders specific for NEC

DM, donor milk; MM, exclusive mother's milk; NEC, necrotizing enterocolitis; PF, preterm formula

Relationship between Cognitive Development and Breast Milk Feeding in Preterm Infants

Background

Many studies have examined the relationship between breastfeeding and cognitive development. Results have been conflicting. Most of the studies were observational in design. Many of them did not have a clear definition of breastfeeding or breast milk exposure. Different cognitive assessment tools were used. Outcomes were measured anywhere from less than 2 years of age to adulthood.

Three systematic reviews from 1999 to 2002 have tried to either establish methodological standards to assess the observational studies or adjust for covariates in pooled analysis (see discussion of confounders in section on cognitive outcomes in term infants). Since the last systematic review by Jain et al. in 2002,⁵⁸ there have been eight cohort studies on preterm infants that examined the relationship of breast milk feeding to some aspects of cognitive development.

A note of caution is in order here. The Mental Developmental Index of the Bayley Scales of Infant Development is widely used to assess the cognitive ability of young children in the studies reviewed in this report. One must keep in mind that the primary purpose of the Bayley Scales is to identify children who may be at risk from developmental delay; it was not the primary purpose to use the results of the Bayley Scales to predict IQ at a later age. Even though some recent studies on preterm infants have shown some predictive ability of the Bayley Scales, the ability to do so is imperfect.^{132,133} Comorbidities (e.g., neurological impairment, extremely low birth weight, other neonatal illnesses), early intervention, environmental, and socioeconomic factors are some of the additional important variables that could affect the prediction of future cognitive function.

Published Systematic Review/Meta-Analysis (Please Refer To Table 8 In Part I)

Studies Identified after the Published Systematic Review/Meta-Analysis (Table 25)

Since 2002, five prospective cohort studies,¹³⁴⁻¹³⁹ two nested case control studies,^{140,141} and one secondary data analysis of a previous randomized controlled trial on supplemental arachidonic and docosahexaenoic acid¹⁴² reported on the relationship between breast milk feeding and some aspects of cognitive development in preterm infants. Sample size of the studies ranged from 39 to 1,035. Five of the studies were of moderate methodological quality,^{135,137,140-142} and three studies were of poor methodological quality.^{134,136,138,139} Each study was graded within its own study design stratum and only with respect to the data on the relationship of breast milk feeding and cognitive development.

Gestational age of the infants ranged from 26 to 33 weeks. Except for one study that specifically enrolled children who had cerebral ultrasound abnormalities, including echodensity, echolucency, and/or ventriculomegaly,¹⁴⁰ the rest of the studies excluded infants with severe congenital abnormalities. Some also excluded infants with perinatal asphyxia and sensorineural abnormalities.

One study reported that half of the subjects had received breast milk exclusively.¹³⁷ Most studies provided information on the amount of breast milk intake and whether the milk had added cow-milk based fortifiers or not while the subjects were in the neonatal wards, but the information on breast milk intake was less informative after discharge to home. Three studies reported the proportion of subjects who breastfed for more than 6 to 7 months; they were 20 percent,¹⁴⁰ 27 percent,¹³⁷ and 29 percent,¹⁴¹ respectively.

Bayley Mental Development Index Scale (MDI) was the cognitive assessment tool for subjects up to 2 years of age. Wechsler Preschool and Primary Scale of Intelligence (WPPSI-R) was used in subjects under 7 years of age. Weschler Intelligence Scale for Children (WISC-R) was used in 7 years and 11 years old. Kaufman Assessment Battery for children (test for overall intellectual function) and Peabody Picture Vocabulary Test (PPVT-R) were administered to 6 to 8 year old children in one study.¹⁴⁰

Prospective cohort. Elgen 2003 prospectively studied 130 low birth weight children at 5 years and 11 years of age using the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-R) and the Weschler Intelligence Scale for Children (WISC-R), respectively.¹³⁵ Twenty-seven percent of them received less than 30 percent breast milk in the neonatal ward. After adjustment for parental confounding (paternal and maternal education), breast milk was no longer a statistically significant predictor of IQ.

Pinelli 2003 prospectively studied 128 infants with birth weight <1,500 g at 6 months and 12 months corrected age using Bayley Mental Development Index Scale (MDI).¹³⁷ Fifty percent of them received breast milk exclusively in the neonatal ward. Twenty-seven percent of infants were breastfeeding at 6 months. After adjustment for sex, SES, birth weight and maternal age, no statistically significant difference in MDI at 12 months was found between the predominantly breastfed group and the predominantly formula-fed group.

The other three lesser quality studies suffered from selection biases (e.g., convenience sample), lack of adjustment for potential confounders (e.g., lack of adjustment for maternal intelligence), small sample sizes, or had issues with incomplete reporting.

One prospective study (in two publications) with post hoc comparisons reported significant differences ($P < 0.05$) in Bayley MDI at 6 months between substantial breast milk feeding group (> 75 percent of nutrition as breast milk) and intermediate (25-75 percent) or minimal feeding groups (< 25 percent) (94.2 ± 8.8 vs. 91.7 ± 7.2 vs. 90.5 ± 8.5 , respectively).^{134,136} Maternal education or SES were not adjusted for in these results.

In a cohort study of subjects from a convenience sample (29 in breast milk group; 10 in formula group), a regression analysis showed an association between the amount of milk infants received in the special care nursery and the Bayley MDI at 7 months ($r=0.4$, $R^2=0.1$, $P < 0.05$) and 12 months ($r= 0.4$, $R^2=0.2$, $P < 0.025$).¹³⁹

In a prospective study of 775 preterm infants who were fed breast milk and 260 preterm infants who were not fed breast milk, the adjusted Bayley MDI at 18 to 22 months of age was 79.9 ± 18 (SD) in the breast milk group versus 75.8 ± 16 (SD) in the no breast milk group ($P = 0.07$).¹³⁸ The result was adjusted for maternal (education, age, marital status) and perinatal (sex, gestational age, oxygen need, birth weight, and illnesses) factors. Mothers in the breast milk group were more likely to have private health insurance, be white and married, and have a college degree. Mothers who had low household income, higher parity, or were black were less likely to provide breast milk feeds. It was unclear if the results were actually adjusted for household income or not. Further analysis of breast milk intake by quintile relative to the no

breast milk group showed that there was a 13.1 point difference in MDI between $\leq 20^{\text{th}}$ quintile and $>80^{\text{th}}$ quintile ($P < 0.0044$).

Re-analysis of a previous RCT on supplemental arachidonic and docosahexaenoic acid. O'Connor 2003 re-analyzed data on 463 subjects from a randomized controlled trial on supplemental formulas in infants less than 33 weeks gestation.¹⁴² Bayley MDI was evaluated at 12 months corrected age. There were no differences in the Bayley MDI among feeding groups. After controlling for home environment and maternal intelligence, there was a significant positive association between duration of breastfeeding and the Bayley MDI at 12 months corrected age in a "full" statistical model ($P = 0.03$, adjusted for large number of preplanned covariates for developmental outcomes), but not in a "reduced" statistical model ($P = 0.07$, adjusted only for those preplanned covariates with a P value of < 0.15).¹⁴²

Nested case-control. Smith 2003 studied 119 preterm subjects with cerebral ultrasound abnormalities and 320 subjects (presumably without ultrasound abnormalities) matched for gestational age in a nested case-control design.¹⁴⁰ The Kaufman Assessment Battery for children were administered to these 6 to 8 years old subjects. In the regression model that included adjustment for maternal verbal ability, home environment, and composite socioeconomic status, the advantage in overall intellectual function associated with direct breastfeeding was 3.6 points (95% CI -0.3 to 7.5; outcome measure was standardized with a mean of 100 points and a standard deviation of 15 points).

Horwood 2001 studied 280 subjects with very low birth weight at 7 years using WISC-R.¹⁴¹ After adjustment for perinatal (sex, gestation, birth weight, multiple birth, Apgar score), socio-demographic (family income, single/two parent family, child ethnicity), and maternal factors (age, education, smoking), there remained a significant association between duration of receipt of breast milk and verbal IQ, with a 6 point advantage for infants who received breast milk for ≥ 8 months compared with no breastfeeding ($P < 0.001$).

Conclusion

No definitive conclusion can be made regarding the relationship between breast milk feeding and cognitive development in preterm infants. One meta-analysis reported a five points advantage in standardized mean score and one systematic review identified one primary study that reported an eight points advantage in IQ in preterm or low birth weight infants who received breast milk feeding. In three of four primary studies of moderate quality that controlled for either maternal education or maternal intelligence, the advantage from breastfeeding was reduced to a statistically non-significant level after adjustment; the fourth study reported a positive association between duration of breastfeeding and the Bayley MDI at 12 months after controlling for maternal intelligence and home environment.

The roles of maternal intelligence and home environment should be accounted for in future studies on breastfeeding and cognitive development. Keeping in mind that cognitive function measured at an early age is not necessarily predictive of later cognitive ability, one should also consider carefully the timing and the selection of appropriate testing instrument in future studies. In addition, clear subject selection criteria, controlling for differences in early complications of prematurity and its relation to receiving breast milk, accounting for subjects lost to follow up, clear distinction between direct breastfeeding and bottle/gavage feeding of breast milk, collect data on breast milk fortifiers or supplemental preterm formulas, and better data collection on breast milk feeding after discharge from the neonatal units will improve the quality of these studies.

Table 25. Summary of primary studies on the relationship of breast milk feeding and cognitive development in preterm infants

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment tools	Confounders adjusted for	Outcomes	Quality and limitations
<i>Prospective cohort</i>						
Elgen 2003 Norway 130 Low birth weight (<2000 g)	Cohort assessed at 5 yr and 11 yr; excluded cerebral palsy, blindness, deafness, multiple malformations, chromosomal abnormalities	27% received <30% breast milk in neonatal ward	Weschler Preschool and Primary Scale of Intelligence- R at 5 yr; Weschler Intelligence Scale for Children-R (WISC-R) at 11 yr	Breast milk, birth weight, paternal and maternal education, chorioamnionitis, gestational age, length of oxygen treatment, sex, 5 minute Apgar, smoking during pregnancy, intrapartum stress, cerebral hemorrhage in subjects with birth weight <1500 g	Unadjusted regression: lack of breast milk was associated with a mean reduction in IQ of 5.8 points (95% CI -11 to -1). After adjustment for parental confounding, breast milk was no longer a statistically significant predictor of IQ.	B Lack of detailed information on breast milk exposure
Pinelli 2003 Canada 148 <1500 g 128 breastfeeding; 20 formula- fed infants were recruited as controls	Cohort from an RCT on conventional breastfeeding support versus supplementary structured breastfeeding counseling; excluded multiple births, severe congenital abnormalities, infants of non- English speaking parents	50% exclusively; some received fortification per explicit criteria; 27% still breastfeeding at 6 mo	Bayley MDI at 6 mo and 12 mo	Sex, SES, birth weight, maternal age	64/128 infants received breast milk exclusively; Adjusted MDI at 6 mo (corrected age) for >80% breastmilk group: 93 (SD 16); for <80% breastmilk group: 94 (SD 15); Adjusted MDI at 12 mo (corrected age) for >80% breastmilk group: 92 (SD 15); for <80% breastmilk group: 91 (SD 12); P>0.05 for all comparisons	B Cognitive testing at a young age

Table 25. Continued

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment tools	Confounders adjusted for	Outcomes	Quality and limitations
Vohr 2006 US 1,035 Preterm with mean gestation of 26.5 wk and birth weight of ~790 g	Enrolled prospectively in the Glutamine Trial at 15 sites of Neonatal Research Network; survivors who received some breast milk and had cognitive testing	Breast milk group– received some breast milk during hospitalization; total volume received calculated per day, values were interpolated for days of the week in which the data were not collected	Bayley Mental Development Index (MDI) at 18 to 22 mos corrected age	marital status, maternal education, age, race/ethnicity, infant's gestational age, gender, culture positive sepsis, grade 3 to 4 intraventricular hemorrhage, oxygen at 36 wks gestational age, necrotizing enterocolitis, and weight <10 th %tile	There were 775 infants in the breast milk group and 260 in the no breast milk group. 95 subjects' Bayley could not be administered successfully; these data were not included. Mothers of infants who were seen at followup were more likely to have received prenatal care than mothers of infants who were not seen. There were no differences in infant characteristics. Information on breast milk feeding was not collected after discharge from hospital. 30.6% of infants in the breast milk group were still receiving breast milk. Mothers in the breast milk group were more likely to be white, married, have a college degree and have private health insurance. Mothers who had low household income, higher parity or were black were less likely to provide breast milk feeds. Adjusted Bayley MDI 79.9 ± 18 in BM 75.8 ± 16 in noBM P=0.0709 MDI <85* 421 (58.1%) in BM 168 (70.9%) in noBM P=0.0355 Multiple regression with adjusted factors=0.53 for MDI, P=0.0002 Breast milk intake was analyzed by quintile relative to the no breast milk group; for MDI, there was a 13.1 point difference between ≤ 20th quintile and >80th quintile (P<0.0044).	C Unclear if family income was adjusted for or not; cognitive test at young age; *proportion of subjects with MDI <85 could not have been adjusted

Table 25. Continued

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment tools	Confounders adjusted for	Outcomes	Quality and limitations
Eidelman, 2004 Feldman, 2003 Israel 86 gestational age 30.5 wk (26-33 wk)	3 groups stratified by amount of breast milk feeding; cognitive assessment at 6 mo corrected age; excluded infants with IVH grade 3 or 4, perinatal asphyxia, metabolic, genetic disease	3 groups: >75% of nutrition as breast milk; 25-75%; and <25%; duration of breast milk feeding: no data	Bayley MDI at 6 mo	Number of breastfeeding episodes was used as a covariate; MANCOVA performed with human milk and infant gender as the between subject factors	Univariate analyses showed group differences on the MDI. Post hoc comparisons reported significant differences between the substantial milk group and the other two groups: 94.2±8.8 vs. 91.7±7.2 vs. 90.5±8.5 (P<0.05)	C Not adjusted for SES, maternal education or intelligence; small sample size; cognitive testing at a young age
Bier 2002 US 39 28.6 wk gestation (range 23-34 wk); 29 in breast milk group; 10 in formula group	convenience sample; excluded infants with mothers who used illicit drugs, had mental illness, HIV infection, and others;	17% of mean volume intake in breast milk group was from premature formula, 9/29 (31%) continued breastfeeding between 3 and 7 mo corrected age	Bayley MDI assessed at 7 mo and 12 mo	maternal result of Peabody Picture Vocabulary Test (PPVT) and days of oxygen	Adjusted for maternal Peabody Picture Vocabulary Test (PPVT) result; mean Bayley MDI at 7 mo, breast milk group = 94 ± 7, in formula group = 90 ± 9 (Difference NS); at 12 mo, breast milk group = 101 ± 11, formula group = 90 ± 9 (P<0.05); regression analysis showed an association between the amount of milk infants received in the special care nursery by gavage and/or bottle and the Bayley MDI at 7 mo (P<0.05) and 12 mo (P<0.025).	C Convenience sample; small sample size, cognitive testing at a young age

Table 25. Continued

Author Year Country Population	Study Description	Breast milk feeding exposure	Assessment tools	Confounders adjusted for	Outcomes	Quality and limitations
<i>Nested case-control</i>						
Smith, 2003 US 119 with ultrasound abnormalities; 320 frequency matched (1:4) were included; gestational age 27.4 wk	Birth subjects from Developmental Epidemiology Network (n=1442); 439 subjects who had undergone neuropsychological testing and for whom parental questionnaires were completed by 9/2002 were included in this study	153 did not receive breast milk; 142 received expressed milk without progressing to direct breastfeeding; 125 received direct breastfeeding (the majority of them first received expressed breast milk); 20% of infants received breast milk > 6 mo.	Kaufman; Peabody Picture Vocabulary Test and Clinical Evaluation of Language Fundamentals; California Children's Verbal Learning Test; at age 6-8 yr	Maternal verbal ability, Home Observation for Measurement of the Environment Inventory, SES, duration of hospitalization	Outcome measures were standardized with a mean of 100 points and standard deviation of 15 points. Breast milk feedings were associated with higher unadjusted test scores for each domain of cognitive function except memory. In the regression model that included social advantage and neonatal morbidity, the adjusted score in overall intellectual function associated with direct breastfeeding was reduced to 3.6 points (95% CI, -0.3 to 7.5).	B 27% of subjects with ultrasound abnormalities; the timing of collection of breastfeeding data was not reported
Horwood, 2001 New Zealand 280 Very low birth weight survivors of a national birth cohort	retrospective recall of breast milk feeding and duration of feeding; assessed at 7 yr; excluded sensorineural disability and incomplete breastfeeding data	Breastfed < 4 mo (n=99) (49%) 4-7 mo (n=46) (22%) ≥ 8 mo (n=59) (29%) Not breastfed (n=76)	WISC-R at 7 yr	Perinatal, socio- demographic (family income, single/two parent family, child ethnicity), and maternal factors (age, education, smoking);	Children who were breastfed ≥ 8 mo had mean verbal IQ 10.2 (SD 0.56) higher and mean performance IQ 6.2 (SD 0.35) higher than children who did not receive breast milk. After adjustment for covariates, there remained a significant association between duration of receipt of breast milk and verbal IQ, with a 6 point advantage for infants who received breast milk for ≥ 8 mo compared with no breastfeeding (P<0.001).	B

Table 25. Continued

Author Year Country	Study Description	Breast milk feeding exposure	Assessment tools	Confounders adjusted for	Outcomes	Quality and limitations
<i>Secondary data analysis of a previous RCT on supplemental arachidonic and docosahexaenoic acid</i>						
O'Connor, 2003 Chile, US, UK 463 <33 wk gestation	Cohort, secondary analysis of a previous RCT on supplemental arachidonic and docosahexaenoic acid; excluded congenital abnormalities that could affect growth and development, major surgery, and others	divided into 4 mutually exclusive groups: predominantly breast milk or formula groups; $\geq$ or < 50% of total energy as breast milk; 43 infants (9%) on predominantly breast milk until term corrected age; 119 (26%) on predominantly formula until term corrected age	Bayley at 12 mo	Maternal intelligence and home environment	No differences in the Bayley MDI were found among feeding groups. There was a positive association between duration of breastfeeding and the Bayley MDI at 12 mo corrected age ($P=0.032$ in full, and $P=0.073$ in reduced statistical models) after controlling for home environment and maternal intelligence.	B Cognitive testing at a young age

Bayley MDI, Bayley Mental Development Index; MANCOVA, Multivariate analysis of variance; NS, Non-significant; PPVT, Peabody Picture Vocabulary Test; SES, socioeconomic status; WISC-R, Weschler Intelligence Scale for Children-R;

Part III. Maternal Outcomes

Relationship between Return to Pre-Pregnancy Weight or Postpartum Weight Change and Breastfeeding

Background

Return to pre-pregnancy weight is desirable since postpartum weight retention is a possible risk factor for obesity and its ensuing medical complications.¹⁴³ The change in weight results from changes in energy metabolism during pregnancy and lactation. This is mediated through the complex neuroendocrine and biochemical stimuli that follow from conception.

Despite the fact that the average weight retention associated with child bearing is modest, estimated at approximately 1.51 kg (s.d.=5.95 kg)¹⁴⁴⁻¹⁴⁶ for some, there is some risk of major weight gain with pregnancy. Ohlin and Rossner 1990 reported changes in body weight that ranged from -12.3 to +26.5 kg from preconception to 1 year postpartum.¹⁴⁴ In various studies, the proportion of women retaining 5 kg or more after 6 months postpartum ranged from 14 to 20%.¹⁴⁶⁻¹⁴⁸ Studies of the impact of physiological and behavioral influences, such as dietary intake, physical activity, and lactation on postpartum weight change reported mixed results. Studies of postpartum weight changes in lactating and non-lactating women also were equivocal within and across populations, with some showing that the length and intensity of breastfeeding were associated with less weight retention after pregnancy, while other studies reported that women who fed their infants formula lost more weight than women who nursed their infants.

Commonly considered confounders in the relationship between return to pre-pregnancy weight or post-partum weight change and breastfeeding were pre-pregnancy weight or BMI, age, educational level, physical activity, parity, smoking status, dieting practice, and ethnicity.

Additional Methodological Comments

To assess the relationship between breastfeeding and return to pre-pregnancy weight and postpartum weight/BMI changes, we relied on Fraser 2003,¹⁴⁹ the only published systematic review of the effect of lactation on maternal body weight. Since Fraser 2003 did not provide any descriptive or summary tables for further analysis, we contacted the authors to obtain copies of their evidence tables. However, the tables made available lacked the appropriate data for additional clarification of the authors' conclusions. Therefore, we decided to perform a primary analysis of all the studies cited in Fraser 2003. We also screened the references in the primary studies to identify additional studies for inclusion in our systematic review. In addition, we also evaluated primary studies that were published after Fraser 2003 and from suggestions by the reviewers of this report. These primary studies had been identified by our general literature search on the intervention of breastfeeding (see Methods chapter). We did not perform a search on maternal weight change by itself without any reference to breastfeeding.

Because of the known methodological problems in the studies of the relationship between breastfeeding and maternal weight,^{15,150,151} we adopted the following inclusion criteria for our review: prospective cohort studies conducted in developed countries which directly compared

weight changes of nonlactating women with that of lactating women and for whom the exclusivity or the amount of breastfeeding was clear. The sample size criterion was at least 50 women per feeding group in the final analyses (e.g., lactating versus nonlactating).

For studies of the relationship between postpartum weight change and breastfeeding, studies need to control for subjects' gestational weight gain or pre-pregnancy weight and have at least 3-months postpartum followup to be included.

Results (Tables 26-27)

A total of 54 potentially relevant articles were retrieved for full-text screening. Forty-five articles were excluded because they were cross-sectional design, non-comparative studies, review articles, sample size less than 50 subjects per group, or had unclear breastfeeding data (exclusivity or amount of breastfeeding not clearly reported).

We included a total of three prospective cohort studies that examined the relationship between exclusive or full breastfeeding and return to pre-pregnancy weight^{145,152,153} and five prospective cohort studies (in six publications) that examined postpartum weight changes in relation to exclusive breastfeeding.^{144,154-157} One study, Linne 2003, was a 15-year followup of the two earlier Ohlin and Rossner studies of 1990 and 1996.

The reporting of breastfeeding data varied across studies. Some studies reported explicit and quantifiable amount of breastfeeding (e.g., providing at least two-thirds of the needed energy intake per kilogram of the infant's weight in breast milk, or infants received 120 mL/day or less of other milk until at least 1 year of age), while others did not quantify the amount of breastfeeding the infants received.

Relationship between breastfeeding and return to pre-pregnancy weight (Table 26)

A total of three prospective cohort studies were identified, involving 4,318, 540, and 95 women, with follow up durations of 3 years, 1 year and 1.5 years postpartum, respectively. All studies were conducted in the United States. All studies enrolled nulliparous and primiparous women between 24 and 40 years of age with normal and above normal pre-pregnancy weights, and all three studies were rated methodological quality grade B.

Overall effect of breastfeeding on return to pre-pregnancy weight or weight retention was negligible. The average weight retention was only within 1 kg range at 1 to 2 years postpartum. The large study of 4,318 nulliparous and primiparous women reported that, compared with women who did not breastfeed, exclusive breastfeeding was associated with a weight gain of approximately 1 kg from pre-pregnancy to 1 to 2 years postpartum, adjusting for age, physical activity and pre-pregnancy BMI.¹⁵² This finding is only statistically significant for nulliparous women who had normal pre-pregnancy weight (BMI < 25) and for those primiparous women who were overweight at baseline (BMI ≥ 25). However, the duration of exclusive breastfeeding was not related to the magnitude of weight change from pre-pregnancy to 1 to 2 year postpartum. The study of 540 parous women reported that breastfeeding at 1 year was significantly associated with less weight retention from first trimester to 1 year postpartum (P = 0.04). The study of 95 nulliparous and primiparous women found that less weight was retained by lactating women than by non-lactating women, and this was statistically significant.¹⁵³ Exclusive breastfeeding was associated with approximately 1 kg weight loss from pre-pregnancy to 1 year postpartum. Bottle-feeding was associated with a weight retention of 2 kg during the same time period. Once lactation was discontinued, slower rates of weight loss were observed. Exclusively breastfeeding

women achieved their pre-pregnancy weights about 6 months earlier than women who exclusively bottle-fed their infants.

Relationship between breastfeeding and postpartum weight changes (Table 27)

Five prospective cohort studies involving a total of 2,097 parous women were identified. Followup durations ranged from 6 months to 15 years. Two studies were conducted in Sweden, two in the United States, and one in Canada. This group of studies measured the post-delivery weight changes in women. Half of the studies did not report women's pre-pregnancy weights. All studies attempted to control for various confounding factors that could influence the relationship between breastfeeding and postpartum weight changes. Five studies were of moderate methodological quality (Grade B), and one of poor methodological quality (Grade C).

The results from the five studies were inconsistent. Among the studies utilized a lactation score to express duration and intensity or exclusivity of breastfeeding, amount of breastfeeding was negatively associated with the postpartum weight change.^{144,157} Among the three studies that examined postpartum weight changes and compared women who exclusively breastfed with women who partially breastfed or exclusively bottle-fed, none of them found a significant relationship between weight loss and breastfeeding among the comparison groups.^{155,156,158}

Conclusion

Based on the results from three prospective cohort studies, we concluded that the overall effect of breastfeeding on return-to-pre-pregnancy weight (weight change from pre-pregnancy or first trimester to 1 to 2 year postpartum) was negligible (less than 1 kg). Results from four prospective cohort studies showed that the effects of breastfeeding on postpartum weight loss were unclear. All seven studies consistently suggested that many other factors have larger effects on weight retention or postpartum weight loss than breastfeeding. Examples of which included annual household income, baseline BMI, ethnicity, gestational weight gain, and energy intake. Undoubtedly, all these factors need to be carefully considered in any future investigation of the relationship between breastfeeding and postpartum weight changes.

Table 26. Summary of prospective cohort studies on the relationship between breastfeeding and returning to pre-pregnancy weights

Author Year Country (N baseline /followup)	Study followup period(s)	Breastfeeding exposure	Comparative feeding groups	Pre- pregnancy weight (kg) / BMI (kg/m ²)	Confounders adjusted for	Results	Quality and limitations
Sichieri 2003 US Nulliparous: 1,538/1,538; Primiparous: 2,810/2,810	3 yr	Exclusive breastfeeding: introduction of daily formula/milk was assumed to represent the end of exclusive breastfeeding period ^a	Never breastfeeding	62 / 82% <25; 18% ≥25	Age, physical activity, BMI, gestational weight gain	After adjusting for confounders, lactation was associated with a weight gain from prepregnancy to 1-2 yr postpartum of approximately 1 kg (statistically significant only for women nulliparous with a baseline BMI <25 (P=0.02) and for those women primiparous with a baseline BMI ≥25 (P=0.04), comparing women who breastfed with women who did not. Duration of exclusive lactation was unrelated to the magnitude of weight change from prepregnancy to 1-2 yr postpartum (P>0.40). Effect of breastfeeding on maternal weight is negligible.	B
Olson, 2003 US 597/540	1 yr postpartum	Breastfeeding after 6 mo was considered to be non-exclusive; a breastfeeding score similar to Ohlin and Rossner's was constructed ^b		ND / 59% ≤26; 41% >26	Age, educational level, smoking status at 1 yr post-partum, annual household income level, pre-pregnancy BMI category, and parity, gestational weight gain	Women who were breastfeeding at 1 yr retained ^c less weight compared with the women who weren't (P = 0.04). Breastfeeding at all other time points and the breastfeeding score were not significantly related to postpartum weight retention. Women at < 185% federal poverty index ratio (PIR) who gained more in pregnancy than Institute of Medicine (IOM) recommendations were 3.73 kg heavier than low income women who gained below or within the IOM guidelines for gestational weight gain (P<0.001).	B

Table 26. continued

Author Year Country (N baseline /followup)	Study followup period(s)	Breastfeeding exposure	Comparative feeding groups	Pre- pregnancy weight (kg) / BMI (kg/m ²)	Confounders adjusted for	Results	Quality and limitations
Janney 1997 US 110/95	0.5, 2, 4, 6, 12, and 18 mo postpartum	Fully breastfeeding was defined as providing at least 2/3 of the needed energy intake per kilogram of the infant's weight in breast milk	Partial breastfeeding; full bottle feeding	59.7 / 22.2	Weight gained during pregnancy, amount of education, parity, type of delivery, season of delivery, marital status, age, smoking status before and during pregnancy, physical activity	Duration of lactation practice was a significant predictor of postpartum weight retention over time (P<0.001). Breastfeeding women achieved their pre-pregnancy weights ~ 6 months earlier than women who bottle-fed their infants. Time returned to prepregnancy weights (mo postpartum): ^d Bottle-feeding only: ~18 Partly breast-feeding at 2 mo, and bottle- feeding thereafter: ~12 Fully breast-feeding for 6 month and bottle- feeding thereafter: ~11 Fully breastfeeding for 6 mo, partly breast- feeding at 12 mo, and bottle-feeding at 18 mo: ~ 10 Marital status is predictor of weight retention over time (0.5-18 mo after parturition) P<0.001. Study showed greater weight retention in unmarried than married women.	B

BMI, Body Mass Index; IOM, Institute of Medicine; ND, not documented; PIR, poverty index ratio

^a Individual breastfeeding data were collected retrospectively in the followup. Women asked to recall their lifetime breastfeeding history.

^b 1 point for each week of coverage of exclusive breastfeeding and 0.5 point for each week of mixed feeding.

^c Weight measured in the first trimester of pregnancy was used to calculate weight retention.

^d Estimated values from figure

Table 27. Summary of prospective cohort studies on the relationship between breastfeeding and postpartum weight/BMI changes

Author Year Country (N baseline /followup)	Study followup period(s)	Breastfeed ing exposure	Comparative feeding groups	Pre- pregnancy weight (kg) / BMI (kg/m ²)	Confounders adjusted for	Results	Quality and limitations
Ohlin 1990; Ohlin 1996 Sweden 2295/1423	1 yr postpartum	Lactation score (scale 0-48): Every month with full lactation was given 4 points, and every month with mixed feeding was given 2 points.		59.6 / 21.5	Weight gain during pregnancy, lactation score, age, pre- pregnancy BMI, parity, weight retention after previous pregnancy, dieting, profession, smoking habits, contraceptive practices, physical activity.	Overall, relationship between the Δ weight and lactation score is statistically significant though weak ($r=-0.09$, $p<0.01$). The weight loss between 2.5 and 6 months postpartum was significantly higher in groups with lactation scores >20 than <20 . However, the total weight loss between 2.5 and 12 months postpartum did not differ significantly between groups. 30% lost weight, 56% gain 0-5 kg, 13% gained 5-10 kg and 1.5% gained ≥ 10 kg 1 year after delivery. Frequency of overweight women ($BMI \geq 23.9$) increased from 13% before pregnancy to 21% 1 year post-partum ($p<0.001$) The coefficients between the Δ weight and lactation score, age, pre-pregnancy weight and parity were very low (multiple $r=0.38$, $p<0.001$). Lactation and age increased the total proportion of explained variance by about 1% each.	B
Linne 2003 Stockholm 1423/563 ^e (SPAWN)	15 yr	Ohlin's lactation score (scale 0- 48)		59.8 / 21.5	As above	Those women who became overweight had lower lactation scores than women who remained normal weight at 15 years follow up (21.7 vs. 24.0, $P < 0.05$).	C High drop out rate

Table 27. continued

Author Year Country (N baseline /followup)	Study followup period(s)	Breastfeeding exposure	Comparative feeding groups	Pre- pregnancy weight (kg) / BMI (kg/m ²)	Confounders adjusted for	Outcomes	Quality and limitations
Walker 2004 US ND/382	6 wk, and 3, 6, and 12 mo postpartum	Full (exclusive) breastfeeding	Partial breastfeeding; full bottle feeding	ND	Ethnicity, maternal education, parity, smoking, physical activity, time of weight measurement, interaction of ethnicity and time, pre- pregnant BMI, gestational weight gain, energy intakes, fat intake, contraception, emotional eating, depressive symptoms	Infant feeding method was not associated with postpartum BMI ($p=0.140$) ^a Ethnicity is associated with BMI ($P=0.001$). In addition, BMI for the ethnic groups is parallel until 3 months postpartum when the African American and Hispanic groups begin increasing in BMI while the White group begins decreasing. At 6 months, the White group continues to decrease, the African American group continues to increase and the Hispanic groups begins to decrease.	B
Haiek 2001 Canada 242/236	6 wk, 3-5, 6-8 mo, and 9 mo postpartum	Predominantly breastfeeding (i.e., exclusive breastfeeding or average daily intake of formula of 4 oz or less)	Mixed-feeding (i.e., average daily intake of formula and breast milk of > 4 oz); predominantly bottle-feeding (i.e., exclusive bottle- feeding or average daily intake of breast milk of 4 oz or less)	22.5	Gestational weight gain, postpartum smoking, infant's solid intake, maternal place of birth	The unadjusted monthly rate of weight change was similar among the three feeding groups for all time periods. For none of the time periods was the rate of weight loss greater in the predominantly breastfeeding group. In the multivariate regression model, neither mixed-feeding nor predominant bottle- feeding was significantly associated with postpartum weight change (breast-feeding group was the reference).	B

Table 27. continued

Author Year Country (N baseline /followup)	Study followup period(s)	Breastfeeding exposure	Comparative feeding groups	Pre- pregnancy weight (kg) / BMI (kg/m ²)	Confounders adjusted for	Outcomes	Quality and limitations
Brewer 1989 US 70/56	3 and 6 mo postpartum	Exclusive breastfeeding	Exclusive formula feeding; mixed breast and formula feeding	ND	Age, parity, pre-pregnancy weight, socioeconomic data, energy intake, energy expenditure exclusive of lactation (physical activity)	Weight loss averaged 1.6, 2.1, and 1.5 kg/mo for exclusive breastfeed, exclusive formula feed, and mixed feed, respectively during the first 3 mo corrected for an initial fluid loss of 2 kg. Exclusive breastfeed resulted in an additional 0.43 kg/mo loss over the second 3 mo (P<0.005) with mixed feed averaging a 0.27 kg loss and nonlactating women experiencing essentially no change. All groups significantly declined in mean maternal weight over the 6-mo postpartum period (p<0.001). No significant differences in total weight loss were observed between the lactating and non- lactating groups.	B

BMI, body mass index.

^a This study is the 15-yr followup of Ohlin 1990;1996. Responders were slightly older, had higher educational attainment, and higher income than non-responders. Nonresponse was more common in non-Nordic citizens.

^b Other variables also included the multivariate model: ethnicity, time, pre-pregnant BMI and other variables (such as parity, gestational weight gain, Cesarean section etc.)

Relationship between Maternal Type 2 Diabetes and Breastfeeding

Background

Studies have shown that lactation has a beneficial effect on glucose and lipid metabolism; and improved pancreatic beta-cell function in women with gestational diabetes.^{159,160} Thus, it is plausible that lactation could reduce the risk of the development of type 2 diabetes.

Commonly considered confounders in the studies of relationship between maternal type 2 diabetes and breastfeeding were parity, body mass index (BMI), diet, physical activity, family history of diabetes, and smoking status.

Published Systematic Review/Meta-Analysis (Table 28)

One systematic review examined the effect of breastfeeding on maternal risk of subsequent diabetes.⁹⁸ The authors identified three studies that evaluated the effects of breastfeeding on glucose tolerance and insulin levels; one study that evaluated the risk factors for recurrent gestational diabetes (GDM); and one study that examined the relationship between lactation and the risk of developing type 2 diabetes in women who had GDM. For this report, we focused only on the one study that examined the relationship between breastfeeding and the risk of developing type 2 diabetes in women with GDM.

Kjos 1998 conducted a retrospective cohort study of 904 Hispanic American women with GDM.¹⁶¹ Kjos 1998 found that the use of progestin-only oral contraceptives was associated with an almost three-fold risk of developing type 2 diabetes compared with the use of combination estrogen-progestin oral contraceptives for breastfeeding Latina women with recent GDM. Since progestin-only oral contraceptive use was invariably associated with breastfeeding in this cohort, the authors then looked for an independent effect of breastfeeding on the risk of developing diabetes among women who initially elected nonhormonal forms of contraception and who were breastfeeding at the time of their initial postpartum examination. The results showed that the risk of developing diabetes in these women was not significantly different from the risk in women who elected nonhormonal contraception but did not breastfeed (unadjusted RR 0.90, 95%CI 0.56 - 1.46; adjusted RR 1.16, 95%CI 0.70 - 1.92). The authors of the systematic review did not assess the quality of this study.

The authors of the systematic review concluded that although no study has ever reported an increased risk of developing diabetes from breastfeeding, a single study did show a potential harm from the use of progestin-only oral contraceptive among breastfeeding Latino women with recent GDM.

Studies Identified after the Published Systematic Review/Meta-Analysis (Table 29)

One prospective, longitudinal cohort study on the association between the duration of lactation and the incidence of type 2 diabetes was identified.¹⁶² The study consisted of two large cohorts in the United States, including participants from the Nurses' Health Study (NHS) and Nurses' Health Study II (NHS II). The Nurses' Health Study (NHS) was initiated in 1976 and

enrolled 121,700 women from 11 states. Participants were between 30 and 55 years of age at baseline, and each woman completed a detailed baseline questionnaire regarding diseases and health related topics. The Nurses' Health Study II (NHS II), begun in 1989, enrolled 116,671 women from 14 states. Participants were between 25 and 42 years of age and completed a similar baseline questionnaire as well as biennial followup questionnaires. This study was graded as high methodological quality (grade A).

The assessments of lactation history and type 2 diabetes were performed longitudinally. Women in the NHS II were also asked to report the diagnosis of GDM on each biennial questionnaire. Cox proportional hazards model was used to calculate the hazard ratio (HR) for type 2 diabetes by lactation history. All models were age-adjusted. Potential confounders including parity, BMI at age 18 years, diet, physical activity, family history of diabetes, and smoking status were included in the multivariate model a priori. Lifetime lactation history among parous women was stratified into six groups: none (reference group), 0 to 3 months, more than 3 to 6 months, more than 6 to 11 months, more than 11 to 23 months, and more than 23 months. Lifetime duration was updated every 2 years. Linear trend was assessed using midpoints of lactation categories. In the analysis of HR per year of lactation, the midpoints of reporting categories to calculate total lifetime lactation were used because this was the closest approximation of the original reported duration.

In the NHS, women who had ever breastfed had a covariate-adjusted HR for type 2 diabetes of 0.97 (95%CI 0.91 - 1.02) compared with women who never breastfed. There was a modest but statistically significant inverse association between duration of lactation and the risk of type 2 diabetes. In the multivariate-adjusted model including current BMI, each additional year of lactation was associated with an HR of 0.96 (95%CI 0.92 - 0.99) for type 2 diabetes.

Among women who had ever breastfed in the NHS II, the covariate-adjusted HR for type 2 diabetes was 0.90 (95%CI 0.77 - 1.04). Each year of lactation was associated with a covariate-adjusted HR of 0.84 (95%CI 0.78 - 0.89). When BMI was added to this model, the HR was 0.88 (95%CI 0.82 - 0.94) for each additional year of lactation.

Women with a history of GDM had a markedly increased risk of type 2 diabetes in the NHS II cohort, with 624 cases per 100,000 person years compared with 118 cases per 100,000 person-years among those without such a history. Lactation had no effect on diabetes risk in the GDM group, with a covariate-adjusted HR of 0.96 (95% CI 0.84 - 1.09) per additional year of lactation.

The effects of exclusive versus total breastfeeding could be compared in the NHS II cohort data. In models controlling for age and parity, each year of lifetime exclusive breastfeeding was associated with an HR for type 2 diabetes of 0.63 (95%CI 0.54 - 0.73), while each year of total breastfeeding was associated with an HR of 0.76 (95%CI 0.71 - 0.81).

Conclusion

Based on the longitudinal study of two large cohorts in the United States with over 150,000 parous women, we conclude that a longer duration of lifetime breastfeeding is associated with a reduced risk of developing type 2 diabetes among parous women who did not have a history of GDM. There was a difference in the risk of developing type 2 diabetes between women with and without GDM in relation to lactation. Compared with women who did not have a history of GDM, women with a history of GDM had a markedly increased risk of type 2 diabetes; and lactation showed no significant relationship with diabetes risk among this group of women. One

must be cautious in interpreting these findings, as they are only generalizable to population with characteristics similar to that of the Nurses' Health cohort.

Table 28. Summary of systematic review/meta-analysis on the relationship between breastfeeding and maternal type 2 diabetes

Author year Population	Study description	Intervention /Comparator	Confounders considered	Results	Quality of SR/MA and limitations
Taylor 2005	SR of 1 retrospective cohort study on the association between breast milk feeding (or lactation) and maternal type 2 diabetes	At their baseline postpartum visits, 461 of the study subjects chose to use an oral contraceptive (OC) and 443 chose a nonhormonal method of contraception. Of those electing to use an OC, 78 were given the progestin-only OC since they were breast-feeding at their baseline postpartum visits and planned to continue breast-feeding. Since progestin-only OC use was invariably associated with breast-feeding in this cohort, the author then looked for an independent effect of breast-feeding on the risk of developing type 2 diabetes among women who initially elected nonhormonal forms of contraception (n=443).	Gestational age at diagnosis of GDM, highest fasting glucose level during the index pregnancy, and glucose area under the curve from the diagnostic OGTT	The risk of developing type 2 diabetes was not significantly different between breastfeeding women and bottle feeding women who chose non-hormonal forms of contraception (adjusted RR: 1.16; 95%CI 0.70-1.92)	C ^a No synthesis of results and unclear how the conclusions were reached

GDM, gestational diabetes; RR, relative risk

^a We did not grade the methodological quality for Kjos 1998 in the systematic review. Taylor 2005 review is also summarized in infant outcome - type 2 DM section.

Table 29. Summary of studies on the relationship between breastfeeding and maternal type 2 diabetes (NIDDM)

Author year Country	Study design	Number of subjects	Mean age at follow- up (year)	Definitions of type 2 diabetes	Breast milk feeding duration (months), parous women only	Adjusted ^a Hazard Ratio (95%CI)	P value for trend	Quality	
Stuebe 2005 US	Prospective longitudinal cohorts (NHS and NHS II studies)	Nurses' Health Study (NHS): 121,700 (83,585 parous women)	NHS: ~52 NHS II: ~35	(1) 1 or more classic symptoms (i.e., excessive thirst, polyuria, weight loss, or hunger) plus either a fasting glucose level of 140 mg/dL (7.8 mmol/L) or more or random plasma glucose level of 200 mg/dL (11.1 mmol/L) or more; (2) at least 2 instances of elevated plasma glucose concentration (fasting glucose ≥140 mg/dL, random plasma glucose ≥200 mg/dL, or oral glucose tolerance test ≥200 mg/dL after 2 hours) on different occasions in the absence of symptoms; or (3) treatment with insulin or an oral hypoglycemic medication.	NHS –	Reference	.02	A	
					None				
					>0 to 3	0.98 (0.91-1.05)			
					>3 to 6	1.03 (0.94-1.13)			
					>6 to 11	0.96 (0.87-1.06)			
					>11 to 23	0.92 (0.84-1.02)			
					>23	0.88 (0.78-1.00)			
					Per year of lactation	0.96 (0.92-0.99)			
					NHS II –	Reference			<.001
					None				
					>0 to 3	1.04 (0.86-1.26)			
					>3 to 6	0.91 (0.73-1.14)			
		>6 to 11	0.87 (0.72-1.06)						
		>11 to 23	0.88 (0.74-1.06)						
		>23	0.67 (0.54-0.84)						
		Per year of lactation	0.88 (0.82-0.94)						
		NHS II, never had GDM –	Reference	<.001					
		None							
		>0 to 3	1.00 (0.81-1.24)						
		>3 to 6	0.80 (0.62-1.04)						
		>6 to 11	0.86 (0.70-1.07)						
		>11 to 23	0.77 (0.62-0.95)						
		>23	0.58 (0.45-0.75)						
		Per year of lactation	0.84 (0.78-0.91)						
NHS II, ever had GDM –	Reference	NS							
None									
>0 to 3	1.18 (0.72-1.93)								
>3 to 6	1.30 (0.75-2.25)								
>6 to 11	0.89 (0.53-1.49)								
>11 to 23	1.19 (0.76-1.87)								
>23	0.97 (0.58-1.61)								
Per year of lactation	0.96 (0.84-1.09)								

GDM, gestational diabetes

^a Confounders adjusted in the model include parity, BMI, diet, physical activity, family history of diabetes, and smoking status.

Relationship between Osteoporosis and Breastfeeding

Background

Osteoporosis is a condition of decreased bone mass. This leads to fragile bones that are at an increased risk for fractures. The World Health Organization (WHO) has established criteria for making the diagnosis of osteoporosis, and for determining levels that predict higher chances of fractures. These criteria are based on comparing bone mineral density (BMD) in a particular patient with those of a healthy 25-year-old female (T-scores). BMD values (T-scores) which fall well below the average for the 25-year-old female (stated statistically as 2.5 standard deviations below the average) are diagnosed as osteoporosis. Although BMD T-scores were based originally on assessment of BMD at the hip by dual-energy X-ray absorptiometry (DXA), they have been applied to define diagnostic thresholds at other skeletal sites and for other techniques. Experts have expressed concern that this approach may not produce comparable data between sites and techniques. Of the various sampling sites, measurements of BMD made at the hip predict hip fracture better than measurements made at other sites, while BMD measurement at the spine predicts spine fracture better than measurements at other sites (consensus.nih.gov/2000/2000Osteoporosis111html.htm).

Calcium and bone metabolism is substantially altered during pregnancy and lactation. The typical daily loss of calcium in breast milk has been estimated to range from 280 to 400 mg, although daily losses as great as 1000 mg calcium have been reported.¹⁶³ Physiologically, during nursing, the body could theoretically meet this demand by increasing the intestinal absorption of calcium, decreasing renal calcium losses, and increasing the resorption of calcium from the maternal skeleton. Bone densities can decrease and increase 3 to 10 percent in the span of a few months in healthy mothers.¹⁶⁴ Prospective cohort studies have reported that lactation is associated with bone mineral loss in the first 6 months to 1 year postpartum, but such loss rebounds overtime.¹⁶⁵⁻¹⁶⁹

Confounders commonly considered in the studies of relationship between fracture risk and breastfeeding were age, hormone replacement therapy, parity and BMI.

Inclusion and Exclusion Criteria

To evaluate the relationship between breastfeeding and the development of osteoporosis, we included all studies that examined the link between breastfeeding and fracture. We also included long-term prospective cohort studies (greater than 1 year of followup duration) that examined the relationship between the duration of breastfeeding and changes in bone mineral densities or bone mineral contents. Articles were excluded if they only used surrogate measures of fracture (e.g., fracture risk score or index) or bone turnover markers.

Results (Tables 30-31)

A total of 44 potential relevant articles were retrieved for full-text screening. Thirty-four articles were excluded due to various reasons (e.g., cross-sectional design, relatively short duration of follow up (< 2 years), studies done in developing countries). We included a total of six case-control studies that examined the risk of fractures in relation to a history of

breastfeeding,¹⁷⁰⁻¹⁷⁵ and four long-term prospective cohort studies that examined the changes of bone mineral densities or bone mineral contents in relation to the duration of breastfeeding.¹⁷⁶⁻¹⁷⁹

Risk of fractures in relation to a history of breastfeeding (Table 30). A total of six case-control studies were identified. There were a total of 1,594 subjects with hip, forearm, or vertebral fractures and 3,523 controls. All subjects were post-menopausal women with an age that ranged from 45 to 103 years old. Three studies were conducted in the United States; one study each was conducted in Australia, Hong Kong, and Sweden. Four studies were rated methodological quality grade B; two was rated grade C.

Incident cases of hip and forearm fractures were identified from hospital or clinical records (with or without radiography confirmations) in five studies. Matched or unmatched general population controls living in the same area were used in three studies, while hospital controls were used in the other two studies. In the remaining one study, all women who were living in three housing blocks were enrolled. Cases of definite vertebral fractures were classified according to the radiological diagnoses and the remaining enrolled subjects without fractures were classified as control subjects. Assessments of breastfeeding history were based on subjects' long-term recalls in all studies. Only one study reported blinded interviewers.

Overall, there was no significant association between a history of breastfeeding and the risk of hip, forearm, or vertebral fractures after adjustment for potential confounders. In four of six studies, parity was considered a potential effect modifier or confounder. Other confounders examined included age, body mass index or weight, hormone replacement therapies, and bilateral oophorectomy. None of the studies provided data on the exclusivity of breastfeeding.

Long-term changes in bone mineral densities or bone mineral contents in relation to the duration of lactation (Table 31). A total of four prospective cohort studies were identified. Followup durations ranged from 2.5 to 12 years. Studies were done in different countries: one studied 113 pre-menopausal parous women in Japan, one studied 169 pre- and peri-menopausal parous women with European heritage living in the United States, one studied 92 pre-menopausal women in Finland, and one studied 121 post-menopausal women in Denmark. One study was of methodological quality grade A, two were of grade B, and one was of grade C, respectively.

Matsushita 2002 examined the effects of multiple pregnancies on BMD of lumbar spine (L2-L4) in 110 parous Japanese women. The outcome was the percent change in BMD, calculated by subtracting the value at the time of the initial pregnancy from the value at the time of the second pregnancy. The results showed that the BMD after the subsequent delivery was significantly higher than the BMD after the initial delivery ($P = 0.001$), with a percent change in BMD of 1.4 percent. Independent determinants of the percent change in BMD were explored by multiple regression analysis. The length of lactation between the deliveries showed no correlation with the percent change in BMD (correlation coefficient = -0.06, $P = 0.702$). Age was the most significant predictor for the percent change in BMD in the model.

Sowers 1992 examined various risk factors for 5-year radial BMD changes in 169 pre- and post-menopausal parous women with European heritage living in the United States. The authors reported that "a recalled history of breastfeeding in parous women did not predict significant differences in BMD level or amount of BMD change".

Uusi-Rasi 2002 examined the relationship between physical activity, calcium intake, and the maintenance of bone mass in 92 non-smoking premenopausal women living in Finland. The effects of total breastfeeding duration on changes in bone mineral contents (BMC) were also

analyzed. According to the multiple stepwise regression analyses, the statistically significant independent predictors for site-specific bone loss were low calcium intake at the baseline and change in body weight both at the proximal femur and at the distal radius sites. In addition, breastfeeding was associated with radial bone loss; the longer the duration of breastfeeding the greater the bone loss (correlation coefficient = -0.34, $P = 0.015$).

Hansen 1991 examined the risk factors for the development of postmenopausal osteoporosis over a 12-year period in 121 postmenopausal women living in Denmark. One hundred eleven of them (92 percent) had one or more pregnancies (mean 2.6) and breastfed their infants for a mean total breastfeeding duration of 13 months. Comparing postmenopausal women who never breastfed their infants with those who did, there was no significant difference in lumbar spine BMD between groups. In addition, there was no significant difference in the annual rate of postmenopausal bone loss between these two groups.

Conclusion

There is no evidence of an association between lifetime breastfeeding duration and osteoporosis. In six case-control studies, there was no significant relationship between a history of lactation and the risk of fractures in postmenopausal women. In two of three moderate or good quality prospective cohort studies using bone mineral density as a surrogate for osteoporosis, lactation does not appear to have an effect on long-term changes in bone mineral densities. The third study found a small decrease in the bone mineral contents in the distal radius with increased duration of breastfeeding, but no significant changes in bone mineral contents in the femoral neck or the trochanter. However, these findings should be interpreted with caution because the feeding history was obtained by maternal recall and data on exclusivity of breastfeeding were not provided. Further investigation with accurate breastfeeding data is warranted.

Table 30. Summary of case-control studies on the relationship between breastfeeding and risk of fracture in post-menopausal women

Table 6a. Summary of case-control studies on the relationship between breastfeeding and risk of fracture in post-menopausal women										
Author year Country	Cases (N)	Controls (N)	Definition of fracture	Mean Age at Dx of fracture (year)	Lactation group	Comparator group	OR (95% CI)		Potential confounders adjusted	Quality and limitations
							Age-adjusted	Full adjusted*		
<i>All post-menopausal women</i>										
Alderman 1986 USA	164	318	Hip fractures based on clinical records	50-74	BF 1-12 mo	Never BF		1.60 (0.91-2.96)	Age, natural logarithm of the duration of estrogen use, and relative weight	B
			BF 13-24 mo				0.40 (0.17-1.10)			
			BF > 24 mo				0.50 (0.21-1.44)			
			Forearm fractures based on clinical records		BF 1-12 mo	Never BF		0.80 (0.45-1.58)		
	180				BF 13-24 mo			0.80 (0.40-1.79)		
					BF > 24 mo			0.80 (0.38-1.87)		
Cumming 1993 Australia	174	137	Hip fractures presenting to the hospitals	65-100	Ever BF	Never BF	0.72 (0.43-1.20)	0.64 (0.30-1.38)	Age, BMI, history of HRT, current use of psychotropic medications, current smoking status, current dairy product consumption, score on mental state, current PA and health status	B
Hoffman 1993 USA	170	173	First hip fracture confirmed by radiography identified from hospital records	50-103	Ever BF	Never BF	0.66 (0.41-1.05)		Age	B
					BF ≤ 12 mo	Never BF	0.67 (0.39-1.13)			
					BF > 12 mo		0.64 (0.32-1.29)			
Chan 1996 Hong Kong	144	163	Definite vertebral fractures by radiography	75	BF < 24 mo	Never BF	0.70 (0.40-1.30)		Age	C Parous women were not separated
					BF ≥ 24 mo		0.60 (0.30-1.00)			

Table 30. Continued

Author year Country	Cases (N)	Controls (N)	Definition of fracture	Mean Age at Dx of fracture (year)	Lactation group	Comparator group	OR (95% CI)		Potential confounders adjusted	Quality and limitations
							Age-adjusted	Full adjusted*		
Kreiger 1982 USA	98	83 ^b 801 ^c	First hip fracture confirmed by radiography identified from hospital records	45-74	BF (12-month increase)			0.50 (0.20-0.90) ^b 0.60 (0.30, 1.0) ^c	Age and BMI, bilateral ovariectomy, and HRT	C Sample size in BF analysis was not described; hospital controls
<i>Parous post-menopausal women</i>										
Michaelsson 2001 Sweden	664	1,848	Hip fractures based on clinical or register records	71	BF 6-10 mo BF 11-16 mo BF > 16 mo BF (3-month increase)	BF 1-5 mo	0.83 (0.66-1.05) 0.84 (0.65-1.08) 0.76 (0.60-0.98) 0.96 (0.94-0.99)	0.87 (0.68-1.11) 0.86 (0.66-1.12) 0.86 (0.67-1.12) 0.98 (0.95-1.01)	Age, HRT, oral contraceptive use, and BMI	B
Alderman 1986 USA	268	211	Hip or forearm fractures based on clinical records	50-74	BF 1-12 mo BF 13-24 mo BF > 24 mo	Never BF		1.20 (0.73-1.95) 0.70 (0.34-1.29) 0.80 (0.38-1.51)	Same as above	B
Cumming 1993 Australia	131	107	Hip fractures presenting to the hospitals	65-100	Ever BF BF 0.5-3 mo ^a BF 3-6 mo ^a BF 6-9 mo ^a BF >9 mo ^a	Never BF Never BF	0.47 (0.22-0.99) 0.56 (0.19-1.68) 0.41 (0.15-1.13) 0.47 (0.19-1.19) 0.28 (0.07-1.18)	0.55 (0.10-2.90) 0.64 (0.13-3.06) 0.79 (0.18-3.51) 0.41 (0.09-1.82) 0.24 (0.04-1.53)	Same as above	B
Hoffman 1993 USA	103	125	First hip fracture confirmed by radiography identified from hospital records	50-103	Ever BF BF ≤ 12 mo BF > 12 mo	Never BF Never BF		0.87 (0.47-1.61) 0.80 (0.42-1.55) 1.08 (0.45-2.60)	Age, and number of live births	B

Dx, diagnosis; BF, breastfeeding; HRT, hormone replacement therapy; PA, physical activity

* Adjusted for potential confounders, including age

^a Average breastfeeding duration per child; ^b trauma controls ; ^c non-trauma controls

Table 31. Summary of prospective cohort studies on the relationship between duration of breastfeeding and the long-term changes of bone mineral density (BMD) or bone mineral content (BMC)

Author, Year Country	Population	Mean baseline age (yr)	Mean duration of followup (yr)	N	Duration of lactation	Δ BMD (g/cm ²) or Δ BMC	p	Quality and limitations
Matsushita, 2001 Japan	Pre-menopausal, parous women	30	2.5	113	Length of lactation between 1 st and 2 nd delivery (mo)	Δ BMD _{spine} : -0.06% ^a	NS	A
Sowers, 1992 USA	Mixed pre- & peri- menopausal, parous women	22-54	5	169	Lactation vs. no lactation	"No significant differences in BMD level or change"	NS	B
Uusi-Rasi, 2002 Finland	Pre-menopausal women	28 (25-30)	4.2 (3.2-5.4)	92	Total duration of breastfeeding ^b	$\% \Delta$ BMC _{femoral neck} ^c $\% \Delta$ BMC _{trochanter} ^c $\% \Delta$ BMC _{distal radius} : $\beta = -0.34$ ^c	NS NS 0.01	B
Hansen, 1991 Denmark	Post-menopausal women	51	12	121	Lactation (87%): mean 13 mo vs. no lactation ^b	Δ BMD _{spine} : 0.04 Δ BMC _{arm} : 0.1% per year	NS NS	C No confounders adjusted; unclear if nonporous women were included

NS, not statistically significant; BMD_{spine}, lumbar spine bone mineral density; BMC_{arm}, forearm bone mineral content

^a Variables in the multiple regression model: interval between the scans (yr), lumbar spine BMD after the initial delivery (g/cm²), bone area (BA) after the initial delivery (cm²), body weight after the initial delivery (kg), age at initial delivery (yr), Δ BA (cm²), body weight changes between the scans (kg) length of lactation between the scans (mo), interval between the time at which breast-feeding was stopped and the next conception (mo); age at next delivery (yr). Dependent variable, Δ BMD%, $r^2=0.174$, $p=0.037$. The percent change in BMD (Δ BMD%) was calculated by subtracting the value at the time of the first pregnancy from the value at the time of the second pregnancy

^b Not clear whether nulliparous women were included in the analyses

^c Multivariate stepwise regression for percent bone loss included the following variables: age, body weight, muscle strength, estimated maximal oxygen uptake and calcium intake, the absolute changes in body weight, muscle strength, estimated maximal oxygen uptake, calcium intake, physical activity, duration of breastfeeding and time from weaning. Only statistical significant variables were kept in the final model.

Relation between Postpartum Depression and Breastfeeding

Background

Postpartum depression is a serious health problem. The prevalence has been estimated at around 13%.¹⁸⁰ It not only affects mother's health, it also affects her ability to care for her infant. Breastfeeding plays a role in affecting an infant's health and in maternal-infant bonding. It is important to understand the nature of the relationship between postpartum depression and the decision to initiate and continue breastfeeding.

Many studies have examined the relationship between breastfeeding and the development of postpartum depression. The results have been quite variable. This may be explained by the lack of a uniform standard in arriving at a diagnosis of postpartum depression. Some studies used questionnaires and some used clinical interviews and different criteria of depression. Many of the studies had relatively small number of subjects. Furthermore, the items commonly used to assess clinical depression like fatigue and sleep problems are to be expected in caring for a newborn. Some of the potential confounders thought to be important in studies of depressive symptoms and feeding practices were marital status, employment status, and whether or not the pregnancy was planned.¹⁸¹

Additional Methodological Comments

We screened the abstracts identified from the MEDLINE general search on breastfeeding in November 2005. Abstracts qualifying for full text retrieval included studies on the relationship between breastfeeding and postpartum depression, psychological disorders, psychiatric illnesses, or mental health issues. We also identified additional articles based on reviews of the bibliographies cited in the relevant retrieved studies from the search. We only included studies that had at least 100 nursing mothers. Qualifying study designs included prospective cohort studies and case-control studies. All methods of assessment of depression were included. Only data pertaining to the relationship of breastfeeding and postpartum depression were extracted from the studies.

Results (Table 32)

Prospective cohort. A total of six prospective cohort studies qualified for inclusion.¹⁸¹⁻¹⁸⁶ There were no case-control studies. The number of women in each study ranged from 113 to 2,375. Three of six studies were rated methodological quality grade B within their respective study design hierarchy and with respect to only the data on the relationship of breastfeeding and postpartum depression. Studies of methodological quality grade C suffered from a combination of incomplete reporting of relevant data, inadequate blinding, lack of or suboptimal adjustment for confounding factors.

Four of six studies did not have specific inclusion criteria based on baseline mental health status. Four studies screened for depression using the Edinburgh Postnatal Depression Scale (EPDS), but the cut off point ranged from 9 to 13 (lowest severity score was zero, highest severity score was 30). Four studies established the diagnosis of depression after clinical interviews. None of the studies provided a clear definition of breastfeeding.

Four prospective cohort studies of moderate methodological quality totaling 4,941 subjects reported postpartum depression rates of 6% to 18%.^{181-183,185} In addition to a history of breastfeeding, all of the studies also considered socio-demographic and obstetric variables as independent predictors of postpartum depression. Assessment of depression by self-reported questionnaires or interviews took place from 1 to 12 months after birth. Except for one study,¹⁸² all of them reported that not breastfeeding or early cessation of breastfeeding was associated with postpartum depression. One study reported that onset of postpartum depression occurred before cessation of breastfeeding in most cases.¹⁸⁵ One study reported that depressed mothers were less likely to continue breastfeeding beyond 2 to 4 months compared with mothers who were not depressed.¹⁸³ In the one study that reported no significant difference in the development of depression in mothers who breastfed versus those who did not breastfeed, women who were breastfeeding at 1 month and were worried about breastfeeding were significantly more likely to become depressed than those who did not worry (RR 3.0, 95%CI 1.041 - 9.216).¹⁸²

Despite the poor methodological quality of the remaining studies, their findings were also consistent with those from studies of moderate quality. One study reported that high EPDS scorers experienced breastfeeding more negatively than the low EPDS scorers (51% versus 16%, $P < 0.0001$).¹⁸⁶ One study reported that mothers who were depressed were less likely to initiate breastfeeding.¹⁸⁴

Conclusion

Studies of moderate quality reported an association between not breastfeeding or short duration of breastfeeding and postpartum depression. More investigation will be needed to determine the nature of this association. It is plausible that postpartum depression led to early cessation of breastfeeding, as opposed to breastfeeding altering the risk of depression. Both effects might occur concurrently. Additional factors that may have a bearing on both postpartum depression and the decision to initiate or terminate breastfeeding should also be sought. Moreover, documentation of baseline mental health status before the initiation of breastfeeding and detailed recording of breastfeeding data will improve the quality of the studies and help understand the nature of the association.

Table 32. Summary of the studies on the relationship between breastfeeding and postpartum depression

Author Year Country Population	Study description	Definition of depression	Confounders adjusted for	Comparators	Results	Quality and limitations
Prospective Cohort						
Warner 1996 UK 2,375 Subjects recruited from postnatal wards prior to discharge	Home interview 6-8 wk after delivery; Edinburgh Postnatal Depression Scale (EPDS) was completed at the interview	Screening: high and low scores using an EPDS threshold of 12/13	Socio-demographic and obstetric variables	Examined "not breastfeeding" as one of the factors in logistic analysis	280/2,375 scored >12 on EPDS; not breastfeeding at 6 wk (OR 1.52, 95%CI 1.12-2.06), unplanned pregnancy (OR 1.44), unemployment in mother (OR 1.56) or in head of household (OR 1.50) were associated with an EPDS >12 in a stepwise logistic analysis.	B No detailed information on breastfeeding
Henderson 2003 Australia 1,745 Subjects recruited from postnatal wards; not under psychological care at the time of recruitment	Self-reported questionnaires were completed at 2, 6, and 12 mo after birth; subjects with high EPDS score were offered a diagnostic interview	Screening: EPDS >12. Diagnosis: DSM IV	Demographic, perinatal, and postnatal factors	Examined "early cessation of breastfeeding" in adjusted hazard ratio	18% developed depression in the 12 mo after birth. Early cessation of breastfeeding was found to be associated with postpartum depression (adjusted hazard ratio 1.25, 95% CI 1.03 to 1.52). Onset of postpartum depression occurred before cessation of breastfeeding in most cases.	B No detailed information on breastfeeding

Table 32. Continued

Author Year Country Population	Study description	Definition of depression	Confounders adjusted for	Comparators	Results	Quality and limitations
Chaudron 2001 US 465 Subjects not depressed at 1 month postpartum	Subjects participated in an interview during the 2 nd trimester, and at 1 and 4 mo after delivery; both the Diagnostic Interview Schedule (DIS) and Center for Epidemiologic Studies Depression Scale (CES-D) were administered at each interview	DSM-III-R criteria; ≥ 16 on the CES-D; and/or receiving antidepressants	Age, depression during pregnancy, postpartum thoughts of death and dying, difficulty falling asleep	Breastfeeding vs. not breastfeeding	6% became depressed between 1 and 4 mo; Women who breastfed their infants were not significantly different in their development of depression from women who did not breastfeed. Of those women who were breastfeeding at 1 mo, women who worried about breastfeeding were significantly more likely to become depressed than those who did not worry (P=0.04)	B No detailed information on breastfeeding

Table 32. Continued

Author Year Country	Study description	Definition of depression	Confounders adjusted for	Comparators	Results	Quality and limitations
Cooper 1993 UK 243 at Oxford; 113 at Cambridge No restriction on mental health status prior to enrollment noted	2 separate cohort studies (Oxford and Cambridge) At Cambridge, prospective subjects were screened at 6 wk postpartum using the EPDS. All high scorers (≥ 13) and other selected participants received full psychiatric assessment between 2 and 3 mo postpartum.	Research Diagnostic Criteria (RDC)	Social class, age, and education	Reported only on "ceased breastfeeding"	At Oxford, 56% of subjects with psychiatric symptoms versus 23% of subjects without psychiatric symptoms had ceased breastfeeding by 8 wk postpartum ($P < 0.01$). In eight, depression preceded cessation of breastfeeding; in two, depression was the subsequent event; and in four, the two events arose contemporaneously. At Cambridge, 56% of subjects with an episode of depression versus 21% of subjects without depression postpartum had ceased breastfeeding by 8 wk ($P < 0.001$). In all cases, the onset of depression preceded the cessation of breastfeeding.	B No detailed information on breastfeeding
Seimyr 2004 Sweden 352 All Swedish speaking pregnant women from six antenatal clinics, with their partners or alone	Administered EPDS at 2 mo before childbirth (I), 2 mo (II) and 12 mo (III) after child birth	Cut-off point of 9/10 on EPDS is used to set the threshold for vulnerability to depression	Did not report actual adjustment, but reported that there was no difference with regards to age, education, parity, and length of marital relationship between low and high-scoring women	Breastfed vs. not breastfed	Fewer high EPDS (I) scorers breastfed compared to the low scorers (82% vs. 94%, $P < 0.02$) Fewer high EPDS (II) scorers breastfed compared to the low scorers (85% vs. 93%, $P < 0.08$) High EPDS scorers experienced breastfeeding more negatively than the women scoring low on EPDS. High EPDS (II) scorers breastfed for a shorter time compared to the low-scoring women (4.6 mo vs. 5.3 mo, $P < 0.04$)	C No adjustment for confounders with respect to breastfeeding

Table 32. Continued

Author Year Country	Study description	Definition of depression	Confounders adjusted for	Comparators	Results	Quality and limitations
Hannah 1992 UK 231 Women who had delivered a live baby	Questionnaires (general questionnaire and EPDS) on 5 th day postpartum and repeat EPDS at 6 weeks postpartum	Postpartum depression defined by EPDS $\geq$ 13 at 6 weeks	Did not report adjustment for breastfeeding, but did report that low mood after a previous delivery was a predictor of post-natal depression	Ever breastfed vs. never breastfed	At 6 wk, 18/26 (69%) women with EPDS $\geq$ 13 vs. 175/200 (88%) women with EPDS < 13 had ever breastfed.	C No adjustment for confounders with respect to breastfeeding

CES-D, Center for Epidemiological Studies Depression Scale; DIS, Diagnostic Interview Schedule; DSM IV, Diagnostic Statistical Manual; EPDS, Edinburgh Postnatal Depression Scale; RDC, Research Diagnostic Criteria

Relationship between Maternal Breast Cancer and Ovarian Cancer and Breastfeeding

Background

Breast cancer is the second most frequently diagnosed and second most deadly cancer among women.¹⁸⁷ Risk factors associated with increased risk of breast cancer include: family history, nulliparity, early menarche, hormone replacement therapy, obesity, and advanced age. Ovarian cancer ranks seventh in the most frequently diagnosed and fourth in the most deadly cancer among women.¹⁸⁷ Risk factors for ovarian cancers are similar to those of breast cancer. While breast and ovarian cancers are closely associated with parity, women with increased parity also have increased lifetime duration of breastfeeding. Therefore, it would be instructive to examine the relationship of breastfeeding and the risk of developing breast or ovarian cancer.

Breast Cancer

Methods

We identified two meta-analyses and one systematic review that evaluated the relationship between breastfeeding and maternal breast cancer.¹⁸⁸⁻¹⁹⁰ We also identified 23 primary studies published since 2001, the cut-off date for literature search used in the latest meta-analysis. Primary studies were screened based on the same inclusion criteria described in the latest meta-analysis. In addition, we elected to include only primary studies conducted in developed countries. Twenty of the 23 primary studies were excluded for the following reasons: studies published in developing countries (5); studies that did not meet the eligibility criteria (e.g., studies that did not provide data on the following: incident invasive breast cancers, reproductive factors, and use of hormonal preparation) (13); duplicate study (1); and review that was not systematically conducted (1). A total of three studies from developed countries published subsequent to the latest meta-analysis were included in the update.¹⁹¹⁻¹⁹³

Published Systematic Reviews and/or Meta-Analyses (Table 33)

The most recent meta-analysis published by the Collaborative Group on Hormonal Factors in breast cancer combined 45 studies published through 2001 and two unpublished studies.¹⁸⁹ The meta-analysis evaluated a total of 50,302 parous women with incident invasive breast cancer and 96,973 controls. Included in the meta-analysis were primary studies that analyzed at least 100 cases of incident invasive breast cancers per study regardless of the menopausal status with additional information on reproductive factors and use of hormonal preparations. Both cohort and case-control studies from developed and developing countries were included. Individual subject data from the primary studies were analyzed for homogeneity across study definitions. The majority of the primary studies did not differentiate between exclusive and partial breastfeeding, and some studies varied in the definition of “ever breastfeeding”. The average age at diagnosis of breast cancer in the studies was 50 years. There was a higher proportion of women with either nulliparity or low parity in the breast cancer group compared with the control group. In addition, lower number of the parous women in the breast cancer group had ever

breastfed their infants compared with the control group. There was a statistically significant reduction in risk of breast cancer by 4.3% (95% CI 2.9-5.8) for each year of breastfeeding. The reduction in the risk of breast cancer with breastfeeding remained unaltered even after stratification for potential confounders such as parity, number of children breastfed, menopausal status, and lifetime duration of breastfeeding. The results were also adjusted for ethnic origin, education, family history of breast cancer, age at menarche, height, weight, body mass index, and use of hormonal contraceptives, alcohol, and tobacco. Decrease in the relative risk of breast cancer associated with each year of breastfeeding remained homogeneous across the studies with regards to developed versus developing countries, age at diagnosis, menopausal status, family history of breast cancer, and study designs. In addition, the decrease in relative risk of breast cancer in parous women, according to breastfeeding history (ever versus never) and number of births, was more pronounced after four or more births. The methodological quality of the meta-analysis was rated grade A.

In the meta-analysis by Bernier 2000 evaluated the relationship of breastfeeding with histologically diagnosed breast cancer in parous women.¹⁸⁸ The meta-analysis combined 25,871 cases and 44,910 controls from 23 primary case-control studies. Criteria for inclusion are studies from developed and developing countries published in English or French languages between 1980 and 1998 that provided usable data for the calculation of odds ratio of breastfeeding and breast cancer risk. The meta-analysis used both fixed and random effects model. There was a small, but statistically significant decreased risk of breast cancer in women who had breastfed their infants compared with women who had not. The decreased risk was further explored by examining the menopausal status at the time of diagnosis of breast cancer and the duration of breastfeeding. Women who were pre-menopausal at the time of diagnosis of breast cancer had a small but statistically significant decreased risk of breast cancer compared with menopausal women. The duration of breastfeeding was divided into categories of 1 to 6 months, 7 to 12 months, and more than 12 months. When these categories of breastfeeding were compared with non-breastfeeding, only those whose lifetime duration of breastfeeding was longer than 12 months (compared to never) had a small but statistically significant reduction in the risk of breast cancer in subgroup analysis. The authors reported that there was no “publication bias”. The methodological quality of the meta-analysis was rated grade B.

Another systematic review included all English language studies published through 1998 and enrolled at least 200 women with breast cancer.¹⁹⁰ In addition, only studies that explicitly adjusted for the number of full-term pregnancies and age at first birth were included. The data included 19,482 cases and 37,627 controls from 24 case-control studies, and 3,857 cases identified from three longitudinal followup studies that comprised of 229,574 subjects. The case-control studies were conducted in hospital- or population-based settings. Studies conducted in developed and developing countries were included. Based on qualitative appraisal, the authors concluded that either there was no relationship between breastfeeding and the risk of development of breast cancer or there was a weak protective effect of ever breastfeeding against the development of breast cancer. The authors did report some reduction in the risk of development of breast cancer in premenopausal women who had breastfed their infants for long duration. The methodological quality of the systematic review was rated grade B.

Studies Published after the Systematic Reviews/Meta-Analyses (Table 34)

We identified three eligible studies from developed countries that were published subsequent to the latest meta-analysis.¹⁹¹⁻¹⁹³ One study was multi-center;¹⁹² the remainders were single center.^{191,193} One study was a prospective cohort of Korean women and evaluated only premenopausal women.¹⁹³ The other two were case-control studies; one evaluated the relationship between breast cancer and breastfeeding among carriers of deleterious BRCA1 and BRCA2 mutations,¹⁹² the other evaluated incident invasive breast cancer patients.¹⁹¹ Two studies included only women whose mean age at diagnosis of breast cancer was in the premenopausal range.^{192,193} All studies reported a statistically significant reduced risk or odds of breast cancer with increased duration of breastfeeding, which varied across the studies. The duration of breastfeeding ranged from 12 to 24 or more months. The methodological quality of the studies ranged from grade B to C.

Conclusion

Results from both meta-analyses concluded that there was a reduction in the risk of breast cancer in women who breastfed their infants. No studies evaluated exclusive breastfeeding. Studies also reported decreased risk or odds of breast cancer in women with a lifetime breastfeeding of more than 12 months. Neither one of the meta-analyses detected any publication bias. In addition, one of the meta-analyses and the systematic review reported decreased odds of breast cancer primarily in premenopausal women. Findings from primary studies published subsequent to the meta-analyses concurred with the findings from the meta-analyses. In conclusion, there is evidence to support the observation that breastfeeding is associated with a reduction in the risk of breast cancer.

Table 33. Summary of systematic reviews/meta-analyses on the relationship between breastfeeding and the risk of breast cancer

Author Year	Number of subjects	Study description	Population	Confounders considered	Intervention /Comparator	Results	Quality for SR/MA
Collaborative Group on Hormonal Factors in Breast Cancer 2002	50,302 cases 96,973 controls	MA of cohort and case-control studies	Incident invasive breast cancer	Ethnic origin, education, family history of breast cancer, age at menarche, height, weight, body mass index, and use of hormonal contraceptives, alcohol, and tobacco	Per 12 mo of breastfeeding	RR reduction of breast cancer 4.3% (99% SE 0.8)	A
					Breastfeeding ever / never	Adjusted RR 0.96 (99% SE 0.2, p=.04)	
					Increasing duration of breastfeeding	Adjusted RR 0.38 (99% SE 0.01, p<.0001)	
Bernier 2000	25,871 cases 44,910 controls	MA of case-control studies	Parous women with histologically diagnosed breast cancer	Age at diagnosis or at full term pregnancy, parity, ethnicity, age at menarche, family history of breast cancer, parity	Breastfeeding ever / never	Summary OR 0.84 (0.78-0.91)	B
					Breastfeeding ever versus never and menopausal at the time of breast cancer diagnosis	Summary OR 0.84 (0.69-1.03)	
					Breastfeeding ever versus never and non menopausal at the time of breast cancer diagnosis	Summary OR 0.81 (0.72-0.91)	
					Breast feeding 0+ to 6 mo versus never	Summary OR 1.00 (0.86-1.16)	
					6+ to 12 mo v never	Summary OR 0.97 (0.86-1.09)	
					12+ mo v never	Summary OR 0.72 (0.65-0.80)	
Lipworth 2000	19,482 cases 37,627 controls	SR of case-control studies	Premenopausal women with breast cancer	Yes, descriptive summary	Breastfeeding ever / never	<i>Qualitative review</i> Author's conclusion "Evidence supporting protective factor of breastfeeding for breast cancer is limited and should be interpreted with caution."	B

MA, meta-analyses; SR, systematic review; RR, relative risk; OR, odds ratio; SE, standard error; NA, not applicable

Table 34. Summary of case-control studies on the relationship between breastfeeding and breast cancer

Author, year Country	Cases (N)	Controls (N)	Definition of disease	Age at diagnosis of disease (year)	Breastfeeding group	Comparator group	Confounders adjusted	Adjusted odds ratio (95% CI)	Quality
Lee 2003 ^a Korea	149		Incident	33-44	1-12 mo	Never (N=161)	Age, oral contraceptives, parity, smoking, exercise and obesity	0.8 (0.7, 1.0) NS	A
	32				13-24 mo			0.7 (0.5, 1.1) NS	
	18				>24 mo			0.6 (0.3, 1.0) P=.04	
Jernstrom 2004 Multi-center	685	685	Invasive BRCA1	39	≤1 yr	Never	Oral contraceptive use and parity	0.89 (0.68, 1.17) NS	B
					>1 yr			0.55 (0.38, 0.80) P=.001	
	280	280	Invasive BRCA2	39	≤1 yr	Never		1.12 (0.73, 1.71) NS	
					>1 yr			0.95 (0.56, 1.59) NS	
Gammon 002 USA	1,508	1,556	Incident; invasive and in-situ	<35 to 85+	<2 mo	Never	Age	0.95 (0.72, 1.24)	C
					2-5 mo			0.82 (0.62, 1.09)	
					6-13 mo			1.06 (0.82, 1.36)	
					14+ mo			0.70 (0.54, 0.92)	

^a Prospective cohort study; Relative risk

Ovarian Cancer

Methods

No meta-analysis was identified on this topic. Eligible designs included prospective cohorts, case-cohort studies or nested case-control studies. Cross-sectional studies were excluded, as described in the Methods section. Eligible studies were conducted in developed countries. We quantified the association between breastfeeding and any type of histopathologically defined maternal ovarian cancer. No specific exclusion criteria were used for the participants in the primary studies.

We recorded or estimated odds ratios for the association between breastfeeding and ovarian cancer for the following comparisons: women who ever breastfed versus never breastfed; women who breastfed less than 12 months (cumulative duration) versus those who never breastfed; and women who breastfed at least 12 months (cumulative duration) versus never. The 12-months cutoff was arbitrary, and was chosen for symmetry with analyses on maternal breast cancer outcomes. At a minimum, studies that were included in the meta-analyses must have been adjusted for parity (or used matching for parity). Adjustment for the use of oral contraceptives (or appropriate matching) was desirable but not mandatory. Studies that reported unadjusted effects only were not included in the meta-analyses, but are reported in the tables and text.

For each study, we estimated the odds ratios for the comparisons of interest when they were not directly reported. This was done for studies that analyzed breastfeeding as an ordinal categorical predictor using cutoffs other than 12 months, provided that mothers who never breastfed were the reference category. In such studies, the odds ratios for the comparisons of interest were estimated by combining odds ratios for different durations of breastfeeding using a random effects model. For example, Riman 2002¹⁹⁴ did not report the odds ratio of ever versus never breastfeeding, but reported odds ratios for specific durations of breastfeeding versus never breastfeeding. Compared with women who never breastfed, the adjusted odds ratio in the Riman study for 1 to 5 months of cumulative breastfeeding was 0.99 (95%CI 0.64 - 1.52); for 6 to 11 months, it was 0.77 (95%CI 0.50 - 1.19) and for at least 12 months of cumulative breastfeeding, it was 0.87 (95%CI 0.56 - 1.35). We can estimate the odds ratio of ever versus never breastfeeding in the Riman 2002 study with a random effects meta-analysis of the aforementioned duration-specific odds ratios (estimated adjusted odds ratio was 0.87, 95%CI 0.68 - 1.12).

Results (Tables 35-38)

We did not identify any prospective studies. We found 15 eligible case-control studies that examined the relationship between breastfeeding and maternal ovarian cancer.¹⁹⁴⁻²¹⁰ In addition, we identified three studies that conducted secondary analyses on population subgroups including white, black, and Jewish women by pooling data from case-control studies.²¹¹⁻²¹³ Population-based controls were used in nine of the 15 studies and hospital-based controls were used in the remaining six. A total of 6,006 subjects with ovarian cancer were studied. The smallest sample included 76 subjects with ovarian cancer matched with 76 hospital-based controls,²⁰⁸ and the largest included 1,028 subjects matched with 2,390 hospital-based controls.¹⁹⁵

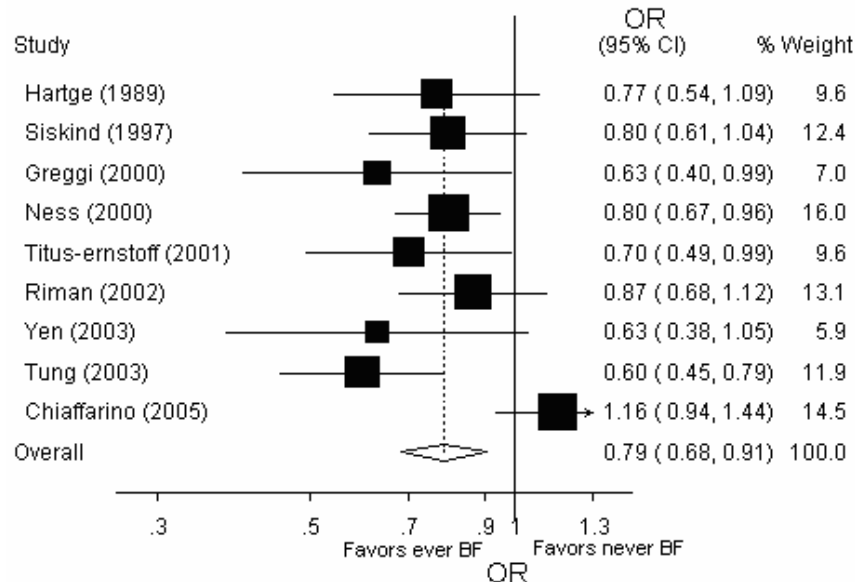
Data were gathered from interviewers using structured questionnaires, and therefore outcome assessment was not blinded. Participant's age ranged from 17 to 79 years. The reporting of information on breastfeeding duration generally did not distinguish between exclusive or partial breastfeeding. Terms commonly used were breastfeeding, lactation, and in one instance, nursing. In 13 out of 15 case-control studies the outcome was histologically confirmed epithelial ovarian cancer. All protocols included women regardless of menopausal status, except for one study that examined peri- and postmenopausal women, and another restricted to women under age 55 years.^{194,198} One primary study included white women only.²⁰² None of the studies verified details of the breastfeeding history by objective data such as hospital or clinic records. Finally, there was wide variability in how breastfeeding duration was categorized in the various studies (Table 35). Because of potential recall bias of breastfeeding history, lack of data verification, and lack of blinding, the majority of the case control studies (n=11) was rated methodological grade B. The remaining four case-control studies were rated grade C. All studies adjusted for parity, and all but one study also adjusted for oral contraceptive use.

Meta-analyses

Ever breastfed versus never breastfed. Nine studies with a total of 4,387 cases and 10,574 controls were included in the comparison of ever versus never breastfeeding (Figure 12). The random effects meta-analysis found an association between breastfeeding and reduced ovarian cancer risk (OR_{adj} 0.79, 95%CI 0.68 – 0.91).^{194,195,197,199-201,204-206,210} There was a statistically significant between-study heterogeneity (P=0.02), which was mainly due to the outlying study of Chiaffarino 2005. Excluding the study yielded very similar estimates without a statistically significant heterogeneity.

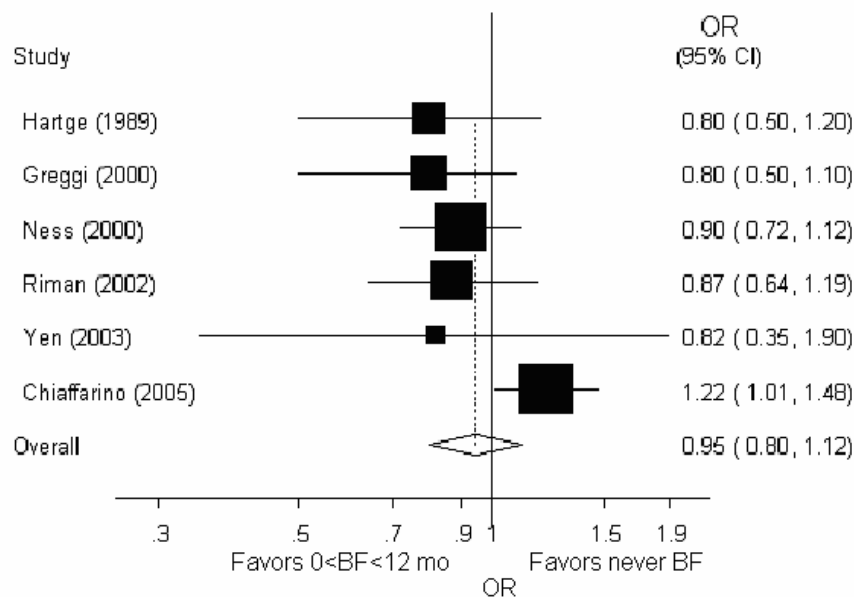
For five of the nine studies we, estimated the odds ratio of ever versus never breastfeeding.^{194,197,199-201,210} If we exclude these five studies from the meta-analysis, a total of 2,582 cases and 4,138 controls remained; and the association between breastfeeding and reduced ovarian cancer risk was no longer statistically significant (OR_{adj} 0.80, 95%CI 0.59 – 1.09)^{195,204-206} Again, there was a statistically significant heterogeneity among the four remaining studies (p<0.01), but the heterogeneity was not readily explained by the characteristics of the primary studies. Excluding the study by Chiaffarino 2005,¹⁹⁵ which was an obvious outlier, the summary odds ratio suggested a slightly stronger (and formally statistically significant) association (0.70 [95%CI, 0.59-0.83]).

Figure 12. Meta-Analysis of case-control studies on the relationship between breastfeeding and maternal ovarian cancer risk: ever breastfed versus never breastfed.



Less than 12 months of cumulative breastfeeding versus never breastfed. Cumulative breastfeeding for less than 12 months was not statistically significantly associated with a decreased risk of ovarian cancer in a meta-analysis of six studies, including 1,911 cases and 5,007 controls in total (OR_{adj} 0.95; 95%CI: 0.80 – 1.12). (Figure 13) The odds ratios for this comparison were estimated for three of the six studies.^{194,195,200,201} There was no statistically significant heterogeneity between the six studies. Two studies (Chiaffarino 2005¹⁹⁵ and Hartge 1989¹⁹⁹) were included in this meta-analysis despite the fact that they used as cutoffs shorter cumulative duration of breastfeeding (10 and 9 months, respectively). Excluding them, the summary of the adjusted odds ratio remained statistically non-significant (0.87, 95%CI 0.74 - 1.02).

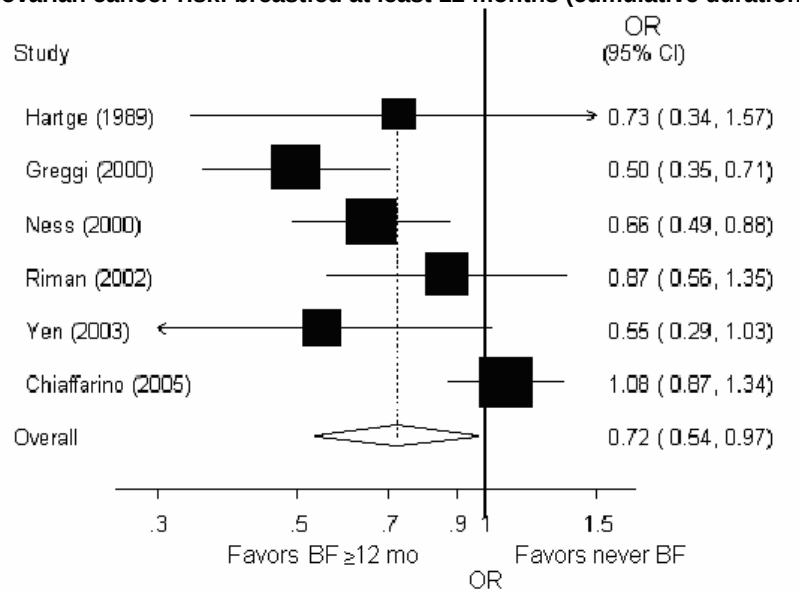
Figure 13. Meta-Analysis of case-control studies on the relationship between breastfeeding and maternal ovarian cancer risk: breastfed less than 12 months (cumulative duration) versus never breastfed



Two studies (Chiaffarino 2005¹⁹⁵ and Hartge 1989¹⁹⁹) were included in this meta-analysis despite the fact that they used as cutoffs shorter cumulative different than 12 months (10 and 9 months, respectively). A sensitivity analysis that excluded them yielded similar inferences (summary of adjusted odds ratios 0.87 [95%CI 0.74, 1.02]).

At least 12 months of cumulative breastfeeding versus never breastfed. Six studies including 1,650 cases and 4,575 controls provided adjusted odds ratios for this comparison,^{194,197,210} or allowed an approximation.^{195,199-201} (Figure 14) Breastfeeding of at least 12 months cumulative duration was associated with 28% lower odds for ovarian cancer (OR_{adj} 0.72, 95%CI 0.54 – 0.97). There was a statistically significant heterogeneity across the six studies. Two studies (Chiaffarino 2005¹⁹⁵ and Hartge 1989¹⁹⁹) were included in this meta-analysis despite the fact that they used as cutoffs shorter cumulative duration of breastfeeding (10 and 9 months, respectively). Excluding them, the summary adjusted odds ratio was even more suggestive of a protective association (0.63 [95%CI 0.50, 0.79]).

Figure 14. Meta-Analysis of case-control studies on the relationship between breastfeeding and maternal ovarian cancer risk: breastfed at least 12 months (cumulative duration) versus never breastfed.



Two studies (Chiaffarino 2005¹⁹⁵ and Hartge 1989¹⁹⁹) were included in this meta-analysis despite the fact that they used as cutoffs shorter cumulative different than 12 months (10 and 9 months, respectively). A sensitivity analysis that excluded them yielded similar inferences (summary of adjusted odds ratios 0.63 [95%CI 0.50, 0.79]).

Subgroup analyses and additional results

Pre- versus post-menopausal women (Table 36). Three studies reported on the risk of ovarian cancer stratified by menopausal status. There were sporadic statistically significant associations between specific categories of breastfeeding duration and reduced risk for ovarian cancer in Sikind 1997²⁰⁴ and Tung 2005.²⁰⁷ Tung 2005 reported a significant dose-response trend between increasing breastfeeding duration and reduction in maternal ovarian cancer risk for post-menopausal women (P value for trend = 0.003). No statistically significant dose-response trends were found in the postmenopausal stratum of Tung 2005 and in both strata in the other two studies. Wynder 1969 reported that in premenopausal women, 10 percent of the cases versus seven percent of the controls breastfed for 12 months or more;²⁰⁹ for postmenopausal women, 24 percent of the cases versus 21 percent of the controls breastfed for 12 months or more. No statistical comparison was reported.

Risk by histologic type (Table 37). Five studies reported the relationship between breastfeeding and the risk of ovarian cancer subgrouped by tumor histology.^{194,195,201,205-207} Comparisons between tumor types included mucinous versus nonmucinous, and invasive versus borderline tumors. Data were limited to allow robust conclusions per histologic type. Moreover, these analyses were subgroup analyses, and therefore, should be viewed as hypothesis forming observations and be interpreted conservatively.

Sporadic associations between specific durations or breastfeeding and reduced risk for histological subtypes were reported in all studies except for Chiaffarino 2005. Breastfeeding was not protective for mucinous or serous cancers consistently across the five studies. Increased duration of breastfeeding was associated with reduced risk for non-mucinous cancers in Tung 2003 (P for trend < 0.001). Titus-Ernstoff 2001 reported an association of breastfeeding with reduced risk for combined endometrioid and clear cell carcinomas, and Riman 2002 reported an

association with reduced risk for clear cell ovarian cancer for a specific duration of cumulative breastfeeding (5 months).

Pooled analysis of subgroups (Table 38). Secondary subgroup analyses were conducted on different racial groups in three studies. John 1993 studied 53 cases of ovarian cancer in black women using data from seven case-control studies,²¹¹ whereas Whittemore 1992 examined data of white women from the same studies and added the cases from five additional studies for a combined population of 1,071 cases. Ever having breastfed was not associated with the risk of development of ovarian cancer in black women but was reported to reduce the risk for white women. Whittemore 1992 divided the sample into hospital and population subjects, the odds ratio for those who breastfed their infants compared with those who did not in the development of ovarian cancer were 0.73 (95%CI 0.51-1.0) and 0.81 (95%CI 0.68-0.95), respectively.²¹² Modugno 2003 analyzed 242 Jewish women from five studies conducted in Israel and the United States.²¹³ This was a cohort study comparing carriers and noncarriers of BRCA1 or BRCA2. The study did not find a difference between carriers and noncarriers of BRCA1 or BRCA2 mutation in the breastfeeding status or breastfeeding duration and the risk of ovarian cancer.

Conclusion

We reviewed 15 case-control studies that examined the relationship between breastfeeding and the risk of ovarian cancer, and performed quantitative syntheses using data from nine studies that adjusted for potential confounders. The overall result from the nine studies suggests an association between breastfeeding and a reduction in the risk of ovarian cancer. Because the reporting in these studies was inconsistent, we needed to estimate the odds ratios in five of nine studies for the meta-analysis. Excluding these five studies results in loss of statistical significance for this association.

There was indirect evidence for a dose-response relationship between breastfeeding and a reduction in the risk of ovarian cancer. Breastfeeding of less than 12 months (cumulative duration) was not statistically significantly associated with a reduction in the risk of ovarian cancer in a meta-analysis of six studies. However, breastfeeding of more than 12 months (cumulative duration) was associated with a reduction in the risk of ovarian cancer, compared with never breastfeeding. We caution that the cutoff of 12 months was arbitrary, and the odds ratios were estimated in half of these studies. Therefore, the interpretation of the postulated dose-response relationship should be done with caution.

Finally, several studies assessed subgroups of pre- and post-menopausal women and ovarian cancer histology. Overall, few sporadic statistically significant associations were identified in some subgroups and for some specific durations of breastfeeding. These findings do not constitute robust evidence.

We conclude that there is some evidence to suggest an association between breastfeeding and a reduction in the risk of maternal ovarian cancer. However, one must be cautious in interpreting this association because it was largely based on estimations of the odds ratios from retrospective studies.

Table 35. Summary of case-control studies on the relationship between breastfeeding and maternal ovarian cancer

Author Year Country	Cases (N)	Controls (N)	Definition of disease	Mean / median age at dx of disease (year)	Breastfeeding group	Comparator group	Confounders adjusted	Adjusted odds ratio (95% CI)	Quality
Chiaffarino 2005 Italy	1,028	2,390	Epithelial ovarian cancer based on histology	56 ^a	Ever BF	Never BF	Age, center, education, parity, oral contraceptive use, family history of ovarian/breast cancer in first degree relatives	1.16 (0.93-1.43)	B
					BF 1-4 mo	Never BF		1.20 (0.91-1.59)	
					BF 5-8 mo			1.24 (0.95-1.62)	
					BF 9-16 mo			1.01 (0.77-1.33)	
					BF ≥17 mo			1.21 (0.85-1.71)	
Riman 2002 Sweden	655	3899	Epithelial ovarian cancer based on histology	Cases 62 Controls 63	BF 1-5 mo	BF <1 mo	Age, BMI, parity, age at menopause, duration of oral contraceptive use, HRT use	0.99 (0.64-1.52)	B
					BF 6-11 mo			0.77 (0.50-1.19)	
					BF ≥12 mo			0.87 (0.56-1.35)	
Siskind 1997 Australia	618	724	Epithelial ovarian cancer based on histology	Cases 57 Controls 56	Ever BF	Never BF	Age, parity, age at 1 st birth, education, oral contraceptive use, smoking history, menopausal status	0.80 (0.61-1.04)	B
					BF 1-6 mo			0.89 (0.65-1.2)	
					BF 7-12 mo			0.68 (0.49-0.94)	
					BF 13-24 mo			0.84 (0.59-1.2)	
					BF 25-36 mo			0.69 (0.38-1.3)	
					> 36 mo			0.77 (0.34-1.8)	
					Reduction risk per month BF			0.99 (0.97-1.0)	

Table 35. Continued

Author Year Country	Cases (N)	Controls (N)	Definition of disease	Mean / median age at dx of disease (year)	Breastfeeding group	Comparator group	Confounders adjusted	Adjusted odds ratio (95% CI)	Quality
Tung 2003; 2005 USA	558	607	Epithelial ovarian cancer based on histology	Cases Mucinous 50 Nonmucinous 55 Controls 56	Ever BF	Never BF	Age, ethnicity, study site, education, oral contraceptive use, tubal ligation	0.6 (0.4-0.7)	B
					BF ≤5 mo			0.6 (0.4-0.9)	
					BF 6-16 mo			0.6 (0.4-0.9)	
					BF >16 mo			0.6 (0.4-0.8)	
Ness 2000 Modugno 2001 USA	531	1190	Epithelial ovarian cancer based on histology	Cases Invasive 53 Borderline 45 Controls 49	BF 1-5 mo	Never BF	Age, number of pregnancies, family history of ovarian cancer, race, oral contraceptive use, tubal ligation, hysterectomy, breastfeeding	0.9 (0.7-1.2)	B
					BF 6-11 mo			0.9 (0.6-1.3)	
					BF 12-23 mo			0.7 (0.5-1.1)	
					BF ≥24 mo			0.6 (0.4-1.0)	
Risch 1994 Canada	450	564	Epithelial ovarian cancer based on histology	Cases 57 Controls 58	Total duration BF of cases (0.51 yr)	Total duration of controls (0.65 yr)	Regression adjusted for 3 age groups, continuous variables age, total duration of oral contraceptive use	0.89 (0.75-1.05)	B
					BF duration/ pregnancy of cases (2.24 mo)	BF duration/ pregnancy of controls (2.72 mo)		0.87 (0.76-0.99)	

Table 35. Continued

Author Year Country	Cases (N)	Controls (N)	Definition of disease	Mean or median age at dx of disease (year)	Breastfeeding group	Comparator group	Confounders adjusted	Adjusted odds ratio (95% CI)	Quality
Gwinn 1990 USA	436	3833	Epithelial ovarian cancer based on histology	20-54	Ever BF	Never BF	Pregnancy, oral contraceptive use, age, and age-pregnancy interaction	Relative risk 0.6 (0.5-0.8)	B
					BF 1-2 mo			0.6 (0.5-0.9)	
					BF 3-5 mo			0.8	
					BF 6-11 mo			0.8	
					BF 12-23 mo			0.7	
					BF ≥24 mo			0.3	
Titus- Ernstoff 2001 USA	378	417	Epithelial ovarian cancer based on pathology or slides	Cases 20-74 Controls matched within 4 years	Ever BF	Never BF	Age, state, parity	0.7 (0.5-1.0)	B
					BF <3 mo			0.9 (0.6-1.4)	
					BF 3-6 mo			0.7 0.5-1.1)	
					BF >6 mo			0.7 (0.4-1.0)	
					2.4% Reduction risk per month BF				
Greggi 2000 Italy	330	721	Ovarian cancer based on histology	54	BF ≤12 mo	Never BF	Age, education, parity, oral contraceptive use, family history of ovarian cancer	0.8 (1.0-1.1)	B
					BF >12 mo			0.5 (0.4-0.8)	
Hartge 1989 USA	203	257	Epithelial ovarian cancer based on histology	Cases 54 Controls 55	BF 1-9 mo	Never BF	Age, race, parity, difficulty conceiving, oral contraceptives, surgical menopause, HRT, or family history of ovarian cancer as needed	Rate ratio 0.8 (0.5-1.2)	B
					BF 10-18 mo			0.5 (0.2-0.9)	
					BF 19-110 mo			1.1 (0.5-2.6)	

Table 35. Continued

Author Year Country	Cases (N)	Controls (N)	Definition of disease	Mean or median age at dx of disease (year)	Breastfeeding group	Comparator group	Confounders adjusted	Adjusted odds ratio (95% CI)	Quality
Yen 2003 Taiwan	86	369H	Epithelial ovarian cancer based on histology	Cases 47 Controls 44	BF ≤1 yr BF >1 yr	Never BF	Parity	0.82 (0.35-1.9) 0.55 (0.29-1.01)	B
Risch 1983 USA	284	705P	Malignant & borderline malignant epithelial ovarian cancer	Cases 20-74 Controls 21- 75	BF ≥3 mo	BF <3 mo ^b	Age at diagnosis 20-44, nulliparous, no miscarriages, < 1 year total exposure to combined oral contraceptives, non-obese	Relative risk 0.69 (0.50-0.96)	C
Cramer 1983 USA	215	215P	Epithelial ovarian cancer	53	Ever BF	Never BF	None	Crude relative risk 1.11 (0.70-1.76)	C
Wynder 1969 USA	158	300H	Epithelial ovarian cancer based on histology	Cases 52 Control ND	No difference between groups for never BF	---	ND	---	C
West 1966 USA	76	76H	Malignant ovarian cancer based on pathology	Cases 53 Controls 50	Cases (6.6 mo) Controls (5.8 mo) p > 0.3	---	None	---	C

Dx, diagnosis; BF, breastfeeding; Mo, month(s); Yr, year(s)

^a Chiaffarino -dx within past year, Risch -dx within past 18 months^b Nulliparous and parous women

Table 36. Summary of case-control studies on the relationship between breastfeeding and maternal ovarian cancer by menopausal status

Author Year Country	Cases (N)	Controls (N)	Definition of disease	Mean or median age at dx of disease (year)	Breastfeeding group	Comparator group	Confounders adjusted	OR _{adj} (95% CI)	p for trend	Quality
Tung 2003; 2005 USA	Pre- menopausal 217	Pre- menopausal 256	Epithelial ovarian cancer based on histology	Cases 55 Controls 55	BF ≤6 mo	Never BF	Age, ethnicity, study site, education, tubal ligation, HRT, ovulation	0.96 (0.49-1.85)	0.003	B
		BF 6-12 mo			0.81 (0.43-1.54)					
		BF >12 mo			0.46 (0.22-0.97)					
	Post- menopausal 341	Post- menopausal 351			BF ≤6 mo	Never BF		0.64 (0.41-1.02)	0.08	
		BF 6-12 mo			0.60 (0.36-1.00)					
		BF >12 mo			0.69 (0.43-1.12)					
Siskind 1997 Australia	Pre- menopausal 215	Pre- menopausal 264	Epithelial ovarian cancer based on histology	Cases 58 Controls 57	BF 1-6 mo	Never BF	Age, parity, age at 1 st birth, education, oral contraceptive use, smoking history, menopausal status	0.75 (0.46-1.21)	ND	B
		BF 7-12 mo			0.53 (0.31-0.94)					
		BF 13-24 mo			1.03 (0.54-1.95)					
		BF >24 mo			0.29 (0.08-1.04)					
	Post- menopausal 403	Post- menopausal 460			BF 1-6 mo	Never BF		0.98 (0.65-1.47)	ND	
		BF 7-12 mo			0.83 (0.54-1.26)					
		BF 13-24 mo			0.88 (0.56-1.38)					
		BF 25-36 mo			0.93 (0.46-1.88)					
		BF >36 mo	1.27 (0.50-3.2)							
Wynder 1969 USA	Pre- menopausal 55	Pre- menopausal 95	Epithelial ovarian cancer based on histology	Cases 52 Control ND	BF >12 mo 9% Cases 7% Controls	---	ND	---	---	C
	Post- menopausal 95	Post- menopausal 205			BF >12 mo 24% Cases 21% Controls					

Dx, diagnosis; BF, breastfeeding; Mo, month(s); Yr, year(s)

Author Year Country	Cases (N)	Controls (N)	Mean age at dx of disease (year)	Histologic type (n)	Breastfeeding group	Comparato r group	Confounders adjusted	Adjusted odds ratio (95% CI)	p for trend	Quality
Chiaffarino 2005 Italy	1028	2390	Cases 56 Controls 57	Serous (492)	Ever BF	Never BF	Age, center, education, parity, oral contraceptive use, family hx of ovarian/ breast cancer in 1° relatives	1.1 (0.85-1.48)	0.71	B
				Mucinous (81)				1.59 (0.82-3.07)	ND	
Titus- Ernstoff 2001 USA	528	523	Cases 20-74 Controls matched within 4 years	Serous borderline (86)	Ever BF	Never BF	Age, state, parity	0.8 (0.4-1.6)	0.61	B
				Serous invasive (229)				1.0 (0.6-1.4)	0.34	
				Mucinous (83)				0.6 (0.3-1.2)	0.88	
				Endometrioid/clear cell (130)				0.4 (0.2-0.7)	0.04	
Tung 2003 USA	558	607	Cases Mucinous 50 Nonmucinous 55 Controls 56	Mucinous (109)	Ever BF BF ≤5 mo BF 6-16 mo BF >16 mo	Never BF	Age, ethnicity, study site, education, oral contraceptive use, tubal ligation	0.8 (0.5-1.4) 0.6 (0.4-1.4) 0.7 (0.3-1.3) 0.9 (0.8-1.8)	0.99	B
				Nonmucinous (449)	Ever BF BF ≤5 mo BF 6-16 mo BF >16 mo	Never BF		0.8 (0.5-1.4) 0.7 (0.4-0.9) 0.5 (0.3-0.7) 0.4 (0.3-0.7)	0.000 5	
Ness 2000 Modugno 2001 USA	531	1,190	Cases Invasive 53 Borderline 45 Controls 49	Borderline Serous (79) Mucinous (60)	Months BF		Age, number of live births, years of oral contraceptive use, years of non- contraceptive estrogen use and months breastfed, tubal ligation, hysterectomy, family history of ovarian and breast cancer, ethnicity	0.95* (0.92-1.00) 0.99 (0.96-1.02)	ND	B
				Invasive Serous (278) Mucinous (52) Endometrioid (136) Other (150)	Months BF			1.00 (0.99-1.01) 1.01 (0.98-1.03) 0.98 (0.95-1.00) 0.97* (0.94-1.00)	ND	

Table 37. Continued

Author Year Country	Cases (N)	Controls (N)	Mean age at dx of disease (year)	Histologic type (n)	Breastfeeding group	Comparato r group	Confounders adjusted	Adjusted odds ratio (95% CI)	p for trend	Qualit y
Riman 2002 Sweden	655	3,899	Cases 62 Controls 63	Invasive Serous (240)	BF 1-5 mo BF 6-11 mo BF ≥12 mo	BF <1 mo	Age, BMI, parity, age at menopause, duration of oral contraceptive use, HRT use	0.87 (0.50-1.53) 0.61 (0.35-1.09) 0.87 (0.49-1.54)	ND	B
				Mucinous (44)	BF 1-5 mo BF 6-11 mo BF ≥12 mo	BF <1 mo		2.19 (0.49-9.87) 1.75 (0.39-7.87) 0.83 (0.17-4.14)		
				Endometrioid (126)	BF 1-5 mo BF 6-11 mo BF ≥12 mo	BF <1 mo		1.05 (0.47-2.34) 1.10 (0.50-2.46) 1.02 (0.44-2.37)	ND	
				Clear cell (25)	BF 1-5 mo BF 6-11 mo BF ≥12 mo	BF <1 mo		0.54 (0.16-1.87) 0.23 (0.06-0.88) 0.24 (0.06-0.97)		

* p<0.05

Table 38. Summary of pooled-analysis of case-control studies on the relationship between breastfeeding and ovarian cancer by population subgroups

Author Year Country	Population characteristic	Cases (N)	Controls (N)	Definition of disease	Mean age at dx of disease (year)	Breastfeeding group	Comparator group	Confounders adjusted	Adjusted odds ratio (95% CI)	Quality
Whittemore ^a 1992 USA	White	201	1,081	Epithelial ovarian cancer	ND	Ever BF	Never BF	Age, study, parity, oral contraceptive use	0.73 (0.51-1.0)	B
						BF 1-5 mo			0.95 (0.62-1.5)	
						BF 6-11 mo			0.40 (0.20-0.78)	
						BF 12-23 mo			0.70 (0.36-1.4)	
						BF ≥24 mo			0.59 (0.22-1.6)	
		870	4,624			Ever BF	Never BF		0.81 (0.68-0.95)	
						BF 1-5 mo			0.87 (0.72-1.1)	
						BF 6-11 mo			0.74 (0.57-0.96)	
						BF 12-23 mo			0.69 (0.51-0.94)	
						BF ≥24 mo			0.74 (0.49-1.1)	
John ^b 1993 USA	Black	53	259	Epithelial ovarian cancer	Cases Invasive 53 Borderline 37 Controls ND	Ever BF	Never BF	Study, birth year, reference age, parity	0.90 (0.42-1.9)	B
						BF 1-5 mo			1.0 (0.39-2.6)	
						≥6 mo			0.85 (0.36-2.0)	
Modugno ^c 2003 USA, Israel	Jewish	95	147	Epithelial ovarian cancer	Cases BRCA1 51 BRCA2 61 Controls BRCA- 58	BRCA1 Ever BF Duration BF	BRCA-	Age at diagnosis, year of birth, number of live births, oral contraceptive use, history of tubal ligation	1.36 (0.68-2.73)	B
									1.01 (0.98-1.04)	
						BRCA2 Ever BF Duration BF	BRCA-		0.70 (0.28-1.72)	
									1.02 (0.99-1.05)	
						BRCA1/2 Ever BF Duration BF	BRCA-		1.09 (0.61-1.97)	
	1.02 (0.99-1.04)									

H-hospital or clinic controls; P, population controls

^a Sample also included Cramer 1983 and Hartge 1989; ^b Sample also included Ness 2000.^c Cases defined as Jewish women with ovarian cancer and BRCA1 or BRCA2 mutation, controls defined as ovarian cancer cases only.

Other Research

An important area of research that is not systematically reviewed in this report is the use of breastfeeding promotion intervention trial to measure health effects (this topic will be covered in a separate report). The best known of these types of studies is the Promotion of Breastfeeding Intervention Trial (PROBIT) conducted in the Republic of Belarus.¹⁷ This was a cluster randomized controlled trial of 34 maternal hospitals and associated polyclinics with a total of 17,046 mother-infant pairs consisting of full term infants and their healthy mothers who intended to breastfeed. The experimental intervention was modeled on the Baby-Friendly Initiative of the World Health Organization and United Nations Children's Fund, which emphasizes assistance with initiating and maintaining breastfeeding and lactation and postnatal breastfeeding support. The control intervention was continuation of the usual infant feeding practices. Results from the study showed that infants in the intervention arm were more likely to be exclusively breastfed at 3 months (43.3% vs. 6.4%; $P < 0.001$) and at 6 months (7.9% vs. 0.6%; $P = 0.01$), and had a significant reduction in the risk of one or more gastrointestinal infections (9.1% vs. 13.2%; adjusted OR 0.60; 95%CI 0.40-0.91) and of atopic dermatitis (3.3% vs. 6.3%; adjusted OR 0.54; 95%CI 0.31-0.95), but no significant reduction in respiratory tract infection. Secondary observational analysis showed that the group of infants who were exclusively breastfed at least 6 months compared to the group of infants who were breastfed 3 to 6 months had a statistically significant reduced risk of one or more episodes of gastrointestinal infection in the first 12 months of life (RR 0.67; 95%CI 0.46-0.97), which was maintained in a multivariate mixed model controlling for geographic origin, urban versus rural location, maternal education, and number of siblings in the household (adjusted OR = 0.61; 95%CI 0.41-0.93).²¹⁴ The same analysis also reported that there was a very low absolute risk of atopic dermatitis in both feeding groups but no risk reduction in the group that was exclusively breastfed for at least 6 months compared with the group that was exclusively breastfed for 3 to 6 months.

Chapter 4. Discussion

Twenty-three outcomes were analyzed in this report. Approximately 400 articles would needed to be reviewed if only articles with primary data were included; this is a much larger volume of literature than can be feasibly reviewed within the time period of this report. With the availability of many published systematic reviews on breastfeeding, we used this literature as the evidence for a large number of outcomes, supplemented by updates of these systematic reviews with new primary studies. We performed several new systematic reviews on outcomes not previously reported. The existing systematic reviews were conducted over a wide span of time and by diverse groups of investigators; there were large variations in the approach and quality of these reviews.

Even though we have assessed the reporting quality of these systematic reviews (using standards of reporting of systematic reviews of observational studies – MOOSE statement²² and additional parameters that we devised), we cannot reliably know the validity of the reported summary data without knowing the details of the primary studies. A number of systematic reviews reported that inclusion and exclusion of some primary studies were reached by consensus between at least two investigators. Without knowing the details of how those consensus were reached, it would be difficult to replicate the findings in those reviews as it is quite plausible that someone who is not familiar with the details of the consensus might have come up with a different set of studies for inclusion in the review. It should also be stressed that a well-performed systematic review does not necessarily imply that the body of evidence for a particular outcome of interest is of high quality. While some systematic reviews assessed the quality of the individual studies, the methods used varied. Any systematic review is limited by the quality of the primary studies included in the review. Unless the method used to assess the quality of the primary studies is transparent and the details made available for examination, it would be difficult to reliably determine the validity of the conclusions.

In most circumstances, it would be unethical to randomize mother-infant pair into breastfeeding (or breast milk feeding) or not breastfeeding arm in a trial. Therefore, the breastfeeding literature is primarily comprised of observational studies, either cohort or case-control studies. There are a number of potential deficiencies related to the study designs that could limit the internal validity and the generalizability of the findings. Some of these potential deficiencies include (1) misclassification of exposure; (2) confounding from the process of self-selection; (3) residual confounding; and (4) insufficient statistical power.

Misclassification of exposure (breastfeeding status/duration) is likely in the studies reviewed in this report. Most studies relied on mothers' recall for the data on breastfeeding. Recall is prone to error. One study from South Africa reported that at 6 to 9 months post-delivery, 13 percent of mothers could not remember the specific timing when they gave something other than breast milk to their infant. In those mothers who could remember, 57 percent of them overestimated the duration of exclusive breastfeeding by about 8 weeks, and 15 percent underestimated the duration by about 3 weeks.²¹⁵ Misclassification, may bias the effect estimate; particularly if the recall error is nonrandom, such as in studies where cases are more likely to underestimate the amount of breastfeeding than controls (for an example, see Norris and Scott 1996⁹⁰).

In studies where subjects were self-selected, there could be confounding from the process of self-selection (e.g., if subjects who perceive their diseases are due to a lack of breastfeeding were more likely to participate in the study).

Residual confounding is a possibility for all observational studies, because it is difficult (if not impossible) to control for all potential confounding variables in these studies. Although it is possible to control for differences in demographic factors, it may not be possible to control for behavioral factors intrinsic in the desire to breastfeed.

Large sample size is often needed to examine the relationship between breastfeeding and various diseases and health conditions because of the need to adjust for numerous confounders to minimize all the potential biases described earlier. It is impossible to predict how these different limitations may interact to increase or decrease the effect estimate.

Compounding the issue of less-than-ideal study design are the heterogeneity of the breast milk itself and differences in how the feedings of breast milk were defined across different studies. The composition of breast milk varies both within and between individuals.^{216,217} The composition could vary depending on preterm versus term delivery, the maternal diet, maternal body weight, time of day, beginning versus near the end of feed, first few months of lactation versus later lactation, milk volume, and numerous other factors. On the other hand, the composition of formulas has also changed significantly over the last twenty years. For example, contemporary formulas have added ingredients like nucleotides and long-chain polyunsaturated fatty acids that were absent from older formulations. How the heterogeneity both within and between comparators would affect the effect estimate is unclear. Also, studies defined breastfeeding differently. Many studies did not have a category of “exclusive breastfeeding”. In the ones that did have this category, “exclusive” could mean no supplement of any kind including water or it could mean occasional formula supplement is permissible. This mixing in of formula in the “exclusive” breastfeeding group may potentially dilute the true effect of breast milk and bias the results toward the null finding. In addition, no study in this review examined the differences between actually breastfeeding an infant and bottle or gavage feeding an infant with breast milk. How the act of breastfeeding itself plays a role in the different effects measured is unknown.

We have summarized the effects of breastfeeding (or breast milk feeding) on a large number of infant and maternal outcomes. Some of the outcomes are well defined and specific (e.g., childhood acute lymphocytic leukemia, breast cancer); and some are not so well defined and non-specific (e.g., asthma, gastrointestinal infections). When the reported outcome is well defined and specific, it lends confidence that the effect reported is valid for that outcome. When the reported outcome is not well defined, one might have some reservation regarding the validity of the measured effect for that outcome.

For all the above reasons, we find that there is a wide range of quality of evidence for the different outcomes examined in this review.

For severe lower respiratory tract diseases, good quality studies did find a relationship between breastfeeding and a reduction in the risk of hospitalization secondary to lower respiratory tract diseases.

For acute otitis media, the results from our meta-analyses of cohort studies of good and moderate methodological quality showed that breastfeeding was associated with a significant reduction in the risk of acute otitis media. Comparing ever breastfeeding with exclusive bottle-feeding, the pooled adjusted odds ratio of acute otitis media was 0.77 (95%CI 0.64 - 0.91). When comparing exclusive breastfeeding with exclusive bottle-feeding, either for more than 3 or 6 months duration, the pooled odds ratio was 0.50 (95%CI 0.36 - 0.70).

For non-specific gastroenteritis, one systematic review identified three primary studies that controlled for potential confounders. These studies reported that there was a reduction in the risk

of non-specific gastrointestinal infections during the first year of life in breastfed infants from developed countries, although the observed range of risk reduction was wide. However, one recent case-control study of 304 infants (167 cases and 137 controls) from England showed that the infants who were breastfeeding had a reduced risk of diarrhea compared to infants who were not breastfeeding (adjusted OR 0.36, 95% CI 0.18 to 0.74, $P=0.005$). The result was adjusted for age, sex, social class, contact with person in and outside household, and other factors. Also, analysis of nested observational cohorts from the Belarus trial showed that the group of infants who were exclusively breastfed for at least 6 months compared to the group of infants who were breastfed for 3 to 6 months had a statistically significant reduced risk of one or more episodes of gastrointestinal infection in the first 12 months of life (adjusted OR = 0.61; 95%CI 0.41-0.93).²¹⁴ The result was adjusted for geographic origin, urban versus rural location, maternal education, and number of siblings in the household.

For necrotizing enterocolitis (NEC) in preterm infants, our meta-analysis of four RCTs found a marginally statistically significant reduction (5% risk difference) of the NEC risk with breast milk feeding. Taking into account the high case-fatality rate of NEC, we consider this estimate is of meaningful clinical difference. However, One must be cognizant of the clinical heterogeneity underlying these RCTs in interpreting the findings of the meta-analysis. Three of the four RCTs were published in the 1980's. Whether infants born in the early 1980's should be combined with infants born in the late 1990's into a meta-analysis is debatable. Neonatal care has made tremendous strides in the last 20 years. Present day preterm formula milk is vastly different from preterm formula milk 20 years ago. All the studies had patient populations that were quite heterogeneous, gestational age ranged from 23 weeks to more than 33 weeks and birth weight ranged from less than 1000 g to more than 1,600 g. One study included only "healthy" infants, another included both "healthy" and "ill" infants. In addition, the types of breast milk, the methods of feedings, and the times of enrollments into the trials were all different. How the heterogeneity in the studies affected the findings is not clear. In addition, studies examining the issue of NEC were frequently also examining the issue of neonatal sepsis, as it is not possible to have NEC without concomitant sepsis. In future studies, it would be worthwhile to examine the relationship of breast milk exposure and sepsis in preterm infants.

For asthma, our subgroup analysis showed that breastfeeding was associated with a reduced risk in children under 10 years of age with a positive family history. However, this association does not hold true for older children as one publication reported a very large adjusted odds ratio (OR 8.7, 95%CI 3.4 – 22.2) for developing asthma in children 6 to 13 years of age who were exclusively breastfed for at least 4 months and had a positive history of maternal asthma.⁵⁶ The relationship of breastfeeding, maternal history, and long-term outcome of asthma bears further investigation.

For atopic dermatitis, available evidence from one well-performed systematic review on full term infants in developed countries suggest that exclusive breastfeeding for at least 3 months confer a protective advantage in the development of atopic dermatitis in those subjects with a family history of atopy. The systematic review did not make a distinction between atopic dermatitis of infancy (under 2 years of age) and persistent or new atopic dermatitis at older ages. This is important because the diagnosis of atopic dermatitis in patients younger than 2 years of age are sometimes attributed to symptoms of infectious origin and breastfeeding may have a protective effect against infections. But a stratified analysis by different durations of followup showed that the risk reduction was similar in those with less than 2 years compared with more than 2 years of followup.

For cognitive outcome in term infants from developed countries, sibling analysis and prospective studies that controlled specifically for maternal intelligence found little or no evidence to support an association between breastfeeding and cognitive performance in children. Most of the published studies adjusted their analyses for socioeconomic status and maternal education but not specifically for maternal intelligence. For those studies that still reported a significant effect after specific adjustment for maternal intelligence, residual confounding from other factors like different home environments cannot be ruled out.

No definitive conclusion regarding the relationship of breast milk exposure and cognitive development in preterm infants can be drawn at this time. Studies that controlled for maternal intelligence reported conflicting results. In addition to maternal intelligence, comorbidities (e.g., neurological impairment, extremely low birth weight, other neonatal illnesses), early intervention, environmental, and socioeconomic factors should also be controlled for in future investigation of this relationship.

For adult blood pressure, evidence suggests that there is an association between a history of breastfeeding during pregnancy and a small reduction in adult blood pressure, but the clinical or public health implication of this finding is unclear. Furthermore, The association weakened after stratification by study size, suggesting the possibility of bias.

For adult cholesterol, a lack of explicit analysis of potential confounders in the meta-analysis hampered the conclusion drawn from the study. Therefore, the relationship between breastfeeding and adult cholesterol levels cannot be adequately addressed at this time.

For cardiovascular mortality in adults, the meta-analysis was limited by the statistical heterogeneity across studies, apparent outcome modification by differences in gender (and therefore, calls into question the appropriateness of combining outcomes from men and women into a single analysis), and more than 30% of the subjects dropped out in the studies. Because of these reasons, no definitive conclusions can be drawn regarding the relationship between breastfeeding and cardiovascular mortality. Further investigation is warranted.

For Sudden Infant Death Syndrome (SIDS), our meta-analysis included only studies that reported clear definitions of exposure, outcomes, and results adjusted for well-known confounders or risk factors for SIDS. The summary estimate found a statistically significant adjusted odds ratio for an association between breastfeeding and a reduced risk of SIDS (adjusted OR 0.64, 95%CI 0.51 - 0.81). We conclude that there is a relationship between breastfeeding and a reduced risk of SIDS. One must be cautious in interpreting this relationship, however. As this finding stems from analysis of observational studies, this finding cannot prove causality. It is plausible that infants who breastfed and breastfed well are less prone to SIDS because of some yet unclarified neurophysiological reasons, and not because breastfeeding itself directly confers a protective effect. Further investigation is warranted.

For post-neonatal mortality (excluding SIDS), there are insufficient data to characterize the relationship between breastfeeding and post-neonatal infant mortality adequately. Further investigation is warranted.

For childhood leukemia, available evidence suggests that there is an association between breastfeeding and a reduced risk of acute lymphocytic leukemia and acute myelogenous leukemia. Our findings from the meta-analyses of the three case-controlled studies that were graded good or fair quality by one systematic review were consistent with the results from the other meta-analysis, but with smaller effect size and smaller statistical significance. Further evaluation of the biological mechanisms underpinning this relationship while taking into

consideration potential biases can be achieved with more large-scale case-controlled studies utilizing population-based and socioeconomic status-matched controls.

For obesity, evidence from three systematic reviews and meta-analyses suggests that a history of breastfeeding is associated with a reduction in the risk of obesity in later life. However, one must be aware of the possibility of residual confounding in interpreting this association. The pooled adjusted odds ratio of obesity comparing ever breastfed to never breastfed was 0.76 (95%CI 0.67-0.86) in one meta-analysis and 0.93 (95%CI: 0.88–0.99) in the other. The magnitude of effects was reduced when more confounders were adjusted in these analyses.

For type 2 diabetes, based on findings from a high-quality systematic review and meta-analyses of seven studies, early breastfeeding was associated with a lower risk of type 2 diabetes in later life compared with those initially formula-fed. However, only three studies appropriately adjusted for all the important confounders, including birth weight, parental diabetes, socioeconomic status, and individual or maternal body size. Even though these three studies found that adjustment did not alter the crude estimate, we cannot be completely confident that potential confounding by birth weight and maternal factors has been ruled out for the overall pooled estimate. This potentially could exaggerate the magnitude of the association.

For type 1 diabetes, even though there are some data to support that breastfeeding for more than 3 months is associated with a reduced risk of type 1 diabetes, this finding must be interpreted with caution because of the likelihood of recall biases and suboptimal adjustments for potential confounders in the primary studies.

For postpartum depression, studies of moderate quality reported an association between not breastfeeding or short duration of breastfeeding and postpartum depression. It is plausible that postpartum depression led to early cessation of breastfeeding, as opposed to breastfeeding altering the risk of depression. Both effects might occur concurrently. Additional factors that may have a bearing on both postpartum depression and the decision to initiate or terminate breastfeeding should be sought. Documentation of baseline mental health status before the initiation of breastfeeding and detailed recording of breastfeeding data will improve the quality of the studies and help understand the nature of the association.

There is no evidence of an association between lifetime breastfeeding duration and maternal osteoporosis. Lactation does not appear to have an effect on long-term changes in bone mineral densities. However, this conclusion is limited by the fact that the feeding history in the studies was obtained by maternal recall and no data on exclusivity of breastfeeding were available. Further investigation with accurate breastfeeding data is warranted.

For breast cancer, there is good evidence to support the observation that breastfeeding is associated with a reduction in the risk of breast cancer. This association is more likely in those women with increased lifetime months of breastfeeding their infants.

For ovarian cancer, there is some evidence to suggest an association between breastfeeding and a reduction in the risk of maternal ovarian cancer. However, one must be cautious in interpreting this association because it was largely based on estimations of the odds ratios from retrospective studies.

For postpartum weight change, we found that the overall effect of breastfeeding on return-to-pre-pregnancy weight (weight change from pre-pregnancy or first trimester to 1 to 2 year postpartum) was negligible (less than 1 kg), and the effect of breastfeeding on postpartum weight change was unclear. Results from the studies also suggest that many other factors have larger effects on weight retention or postpartum weight loss than breastfeeding. Methodological

challenges in these studies include the accurate measurement of energy balance, adequate control for numerous covariables, and quantifying accurately the exclusivity and the duration of breastfeeding. None of the included studies tackled all of these challenges.

Concerning the risk of maternal type 2 diabetes, a longer duration of lifetime breastfeeding is associated with a reduced risk of developing type 2 diabetes among parous women who did not have a history of gestational diabetes (GDM). There was a difference in the risk of developing type 2 diabetes between women with and without GDM in relation to lactation. Compared with women who did not have a history of GDM, women with a history of GDM had a markedly increased risk of type 2 diabetes; and lactation showed no significant relationship with diabetes risk among this group of women. One must be cautious in interpreting these findings, as they are only generalizable to population with characteristics similar to that of the Nurses' Health cohort.

An important area of research that is not systematically reviewed in this report is the use of breastfeeding promotion intervention trial to measure health effects (this topic is not part of the scope of this report and it will be covered in a separate report). The best known of these types of studies is the previously described Promotion of Breastfeeding Intervention Trial (PROBIT) conducted in the Republic of Belarus.¹⁷ Data from this study provided good evidence that breastfeeding is associated with a reduction in the risk of gastrointestinal infection and atopic dermatitis. Whether results from studies conducted in other countries are applicable to the United States is unclear. In Belarus, mothers often stay in the hospital close to one week post delivery, infant formulas can cost as much as 20% of an average salary, and there is an obligatory prolonged maternity leave (approximately 3 years in most cases). In contrast, in the United States, mothers are often discharged within 48 hours post delivery and formula manufacturers provide rebates to the Special Supplemental Nutrition Program for Women, Infants, and Children (www.wicprogram.org). The factors in Belarus could work in conjunction with the intervention to help promote the increase in the rate of exclusive breastfeeding. On the other hand, one may argue that the results reported in the Belarus study could serve as a best-case scenario in terms of the potential benefits of breastfeeding when optimal promotion and support of breastfeeding are in place. More research in this country along the line of the Belarus study should be considered.

Of note, there were a few individual primary studies on asthma, cardiovascular mortality, and type 1 diabetes that reported increase in risk of those diseases in subjects who had been breastfed. Even though those studies were few in numbers, those findings should not be ignored and further investigation should be done.

Lastly, the outcomes analyzed in this review represent only a portion of all possible health outcomes related to breastfeeding reported by investigators worldwide. To work within the constraints of resources, we relied on the advice from our panel of technical experts in finalizing the list of outcomes included in this review. Thus, some important outcomes (e.g., growth and nutrition) have, by necessity, not been included in this review. Additional systematic reviews germane to those important outcomes would be of value.

Future Research

Assessment of the association between breastfeeding and health outcomes

Observational studies will remain the major source of information in this field. Clear subject selection criteria, adopting a common definition of “exclusive breastfeeding”, reliable collection of feeding data, specific and properly quantifiable outcomes of interest, controlling for important potential confounders including child-specific factors, and blinded assessment of the outcome measures will help immeasurably to improve the quality of these studies. Traditional

retrospective case-control studies, usually used when the disease is rare, are less desirable because of the many caveats noted earlier. Prospective nested case-control studies with blinded assessment of the outcome measures would provide more reliable results.

As have been mentioned previously, it is not possible to eliminate self-selection bias in observational studies because of behavioral or attitudinal factors intrinsic in the desire to breastfeed. Thus, it is worthwhile to study these factors to further understand the reasons for the decision to breastfeed.

Sibling analysis provides a method to control for hereditary and household factors that are important in certain outcomes, provided that those factors are similar for the siblings of interest. Although such analysis may be less susceptible to confounders and effect modifiers that are shared by siblings, one must remember that it is not immune to biases. This method should be used when the appropriate data are available.

There is a large degree of heterogeneity across studies among many of the outcomes. The heterogeneity persisted after adjusting for potential confounders. It might be helpful to study breast milk composition (e.g., oligosaccharides, nucleotides, and others) with respect to the residual heterogeneity. In addition, maternal genetic variations in the production of those factors of interest from breast milk can be studied (for an example, see the discussion by Newburg 2005²¹⁸ concerning the variability of antidiarrheal effect of breastfeeding according to the prevalence of the secretor gene (*fucosyltransferase 2*) and the Lewis gene (*fucosyltransferase 3*) in the study population).

Assessment of the efficacy/effectiveness of breastfeeding promotion interventions

Cluster randomized controlled studies similar to the Belarus trial will provide understanding of the effectiveness of various breastfeeding promotion interventions. Any substantial differences in the degree of breastfeeding between the two groups as a result of the intervention will provide further opportunity to investigate any disparity in health outcomes between the two groups.

References and Included Studies

1. American Academy of Pediatrics Committee on Nutrition. Pediatric Nutrition Handbook. Fifth. 2004.
2. U.S.Department of Health and Human Services. Healthy People 2010: Conference Edition. Washington, DC: U.S. Government Printing Office; 2000.
3. Scariati PD, Grummer-Strawn LM, Fein SB. A longitudinal analysis of infant morbidity and the extent of breastfeeding in the United States. *Pediatrics* 1997;99(6):E5.
4. Gillman MW, Rifas-Shiman SL, Camargo CA, Jr., et al. Risk of overweight among adolescents who were breastfed as infants. *JAMA* 2001;285(19):2461-7.
5. Grummer-Strawn LM, Mei Z. Does breastfeeding protect against pediatric overweight? Analysis of longitudinal data from the Centers for Disease Control and Prevention Pediatric Nutrition Surveillance System. *Pediatrics* 2004;113(2):e81-e86.
6. Oddy WH, Holt PG, Sly PD, et al. Association between breast feeding and asthma in 6 year old children: findings of a prospective birth cohort study. *BMJ* 1999;319(7213):815-9.
7. Gartner LM, Morton J, Lawrence RA, et al. Breastfeeding and the use of human milk. *Pediatrics* 2005;115(2):496-506.
8. Chua S, Arulkumaran S, Lim I, et al. Influence of breastfeeding and nipple stimulation on postpartum uterine activity. *Br J Obstet Gynaecol* 1994;101(9):804-5.
9. Dewey KG, Heinig MJ, Nommsen LA. Maternal weight-loss patterns during prolonged lactation. *Am J Clin Nutr* 1993;58(2):162-6.
10. Newcomb PA, Storer BE, Longnecker MP, et al. Lactation and a reduced risk of premenopausal breast cancer.[see comment]. *N Engl J Med* 1994;330(2):81-7.
11. Rosenblatt KA, Thomas DB. Lactation and the risk of epithelial ovarian cancer. The WHO Collaborative Study of Neoplasia and Steroid Contraceptives. *Int J Epidemiol* 1993;22(2):192-7.
12. Gross SJ. Growth and biochemical response of preterm infants fed human milk or modified infant formula. *N Engl J Med* 1983;308(5):237-41.
13. Tyson JE, Lasky RE, Mize CE, et al. Growth, metabolic response, and development in very-low-birth-weight infants fed banked human milk or enriched formula. I. Neonatal findings. *J Pediatr* 1983;103(1):95-104.
14. Li R, Darling N, Maurice E, et al. Breastfeeding rates in the United States by characteristics of the child, mother, or family: the 2002 National Immunization Survey. *Pediatrics* 2005;115(1):e31-e37.
15. Bauchner H, Leventhal JM, Shapiro ED. Studies of breast-feeding and infections. How good is the evidence? *JAMA* 1986;256(7):887-92.
16. Kramer MS. Does breast feeding help protect against atopic disease? Biology, methodology, and a golden jubilee of controversy. *J Pediatr* 1988;112(2):181-90.
17. Kramer MS, Chalmers B, Hodnett ED, et al. Promotion of Breastfeeding Intervention Trial (PROBIT): a randomized trial in the Republic of Belarus. *JAMA* 2001;285(4):413-20.
18. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7(3):177-88.
19. Higgins JP, Whitehead A, Turner RM, et al. Meta-analysis of continuous outcome data from individual patients. *Stat Med* 2001;20(15):2219-41.
20. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ* 2003;327(7414):557-60.
21. Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999;354(9193):1896-900.
22. Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;283(15):2008-12.

23. Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. *Ann Intern Med* 2001;134(8):663-94.
24. Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001;285(15):1987-91.
25. Auinger P, Lanphear BP, Kalkwarf HJ, et al. Trends in otitis media among children in the United States. *Pediatrics* 2003;112(3 Pt 1):514-20.
26. Bluestone C, Klein J. Otitis media in infants and children. 2nd edition. WB Saunders; 1995.
27. Nylen O, Anderson B, Aniansson G. Reduced frequency of acute otitis media in breast-fed infants. Abstract from the 5th International Congress of Pediatric Oto-Rhino-Laryngology Gent, Belgium, June 5-9. 1990.
28. Uhari M, Mantysaari K, Niemela M. A meta-analytic review of the risk factors for acute otitis media.[see comment]. *Clin Infect Dis* 1996;22(6):1079-83.
29. Duncan B, Ey J, Holberg CJ, et al. Exclusive breast-feeding for at least 4 months protects against otitis media.[see comment]. *Pediatrics* 1993;91(5):867-72.
30. Howie PW, Forsyth JS, Ogston SA, et al. Protective effect of breast feeding against infection. *BMJ* 1990;300(6716):11-6.
31. Teele DW, Klein JO, Rosner B. Epidemiology of otitis media during the first seven years of life in children in greater Boston: a prospective, cohort study. *J Infect Dis* 1989;160(1):83-94.
32. Vernacchio L, Lesko SM, Vezina RM, et al. Racial/ethnic disparities in the diagnosis of otitis media in infancy. *Int J Pediatr Otorhinolaryngol* 2004;68(6):795-804.
33. Duffy LC, Faden H, Wasielewski R, et al. Exclusive breastfeeding protects against bacterial colonization and day care exposure to otitis media. *Pediatrics* 1997;100(4):E7.
34. Alho OP, Laara E, Oja H. Public health impact of various risk factors for acute otitis media in northern Finland. *Am J Epidemiol* 1996;143(11):1149-56.
35. Alho OP, Laara E, Oja H. How should relative risk estimates for acute otitis media in children aged less than 2 years be perceived? *J Clin Epidemiol* 1996;49(1):9-14.
36. Sassen ML, Brand R, Grote JJ. Breast-feeding and acute otitis media. *Am J Otolaryngol* 1994;15(5):351-7.
37. Stenstrom C, Ingvarsson L. Otitis-prone children and controls: a study of possible predisposing factors. 1. Heredity, family background and perinatal period. *Acta Oto-Laryngologica* 1997;117(1):87-93.
38. Schultz Larsen F, Hanifin J. Epidemiology of Atopic Dermatitis. *Immunol Allergy Clin North Am* 2002;22(1):1-24.
39. Saarinen UM, Kajosaari M, Backman A, et al. Prolonged breast-feeding as prophylaxis for atopic disease. *Lancet* 1979;2(8135):163-6.
40. Gdalevich M, Mimouni D, David M, et al. Breast-feeding and the onset of atopic dermatitis in childhood: a systematic review and meta-analysis of prospective studies. *J Am Acad Dermatol* 2001;45(4):520-7.
41. Imhoff B, Morse D, Shiferaw B, et al. Burden of self-reported acute diarrheal illness in FoodNet surveillance areas, 1998-1999. *Clin Infect Dis* 2004;38 Suppl 3:S219-S226.
42. Feachem RG, Koblinsky MA. Interventions for the control of diarrhoeal diseases among young children: promotion of breast-feeding 10045. *Bull World Health Organ* 1984;62(2):271-91.
43. Morrow AL, Ruiz-Palacios GM, Jiang X, et al. Human-milk glycans that inhibit pathogen binding protect breast-feeding infants against infectious diarrhea. *J Nutr* 2005;135(5):1304-7.
44. Newburg DS, Peterson JA, Ruiz-Palacios GM, et al. Role of human-milk lactadherin in protection against symptomatic rotavirus infection. *Lancet* 1998;351(9110):1160-4.
45. Chien PF, Howie PW. Breast milk and the risk of opportunistic infection in infancy in industrialized and non-industrialized settings. [Review] [75 refs]. *Adv Nutr Res* 2001;10:69-104.
46. Quigley MA, Cumberland P, Cowden JM, et al. How protective is breast feeding against diarrhoeal disease in infants in 1990s England? A case-control study. *Arch Dis Child* 2006;91(3):245-50.

47. Shay DK, Holman RC, Newman RD, et al. Bronchiolitis-associated hospitalizations among US children, 1980-1996. *JAMA* 1999;282(15):1440-6.
48. Bachrach VR, Schwarz E, Bachrach LR. Breastfeeding and the risk of hospitalization for respiratory disease in infancy: a meta-analysis. *Arch Pediatr Adolesc Med* 2003;157(3):237-43.
49. Eder W, Ege MJ, von ME. The asthma epidemic. *N Engl J Med* 2006;355(21):2226-35.
50. Strachan DP. Hay fever, hygiene, and household size. *BMJ* 1989;299(6710):1259-60.
51. Braun-Fahrlander C, Riedler J, Herz U, et al. Environmental exposure to endotoxin and its relation to asthma in school-age children. *N Engl J Med* 2002;347(12):869-77.
52. Gdalevich M, Mimouni D, Mimouni M. Breast-feeding and the risk of bronchial asthma in childhood: a systematic review with meta-analysis of prospective studies. *J Pediatr* 2001;139(2):261-6.
53. Burgess SW, Dakin CJ, O'Callaghan MJ. Breastfeeding does not increase the risk of asthma at 14 years. *Pediatrics* 2006;117(4):e787-e792.
54. Kull I, Almqvist C, Lilja G, et al. Breast-feeding reduces the risk of asthma during the first 4 years of life.[see comment]. *J Allergy Clin Immunol* 114(4):755-60, 2004;114(4):755-60.
55. Sears MR, Greene JM, Willan AR, et al. Long-term relation between breastfeeding and development of atopy and asthma in children and young adults: a longitudinal study.[see comment]. *Lancet* 2002;360(9337):901-7.
56. Wright AL, Holberg CJ, Taussig LM, et al. Factors influencing the relation of infant feeding to asthma and recurrent wheeze in childhood.[see comment]. *Thorax* 2001;56(3):192-7.
57. Drane DL, Logemann JA. A critical evaluation of the evidence on the association between type of infant feeding and cognitive development. *Paediatr Perinat Epidemiol* 2000;14(4):349-56.
58. Jain A, Concato J, Leventhal JM. How good is the evidence linking breastfeeding and intelligence? *Pediatrics* 2002;109(6):1044-53.
59. Johnson DL, Swank PR, Howie VM, et al. Breast feeding and children's intelligence. *Psychol Rep* 1996;79(3 Part 2):1179-85.
60. Wigg NR, Tong S, McMichael AJ, et al. Does breastfeeding at six months predict cognitive development?[see comment]. *Aust N Z J Public Health* 1998;22(2):232-6.
61. Anderson JW, Johnstone BM, Remley DT. Breast-feeding and cognitive development: a meta-analysis. *Am J Clin Nutr* 1999;70(4):525-35.
62. Der G, Batty D, Deary IJ. Effect of breast feeding on intelligence in children: prospective study, sibling pairs analysis, and meta-analysis. *BMJ* 2006.
63. Agostoni C, Marangoni F, Giovannini M, et al. Prolonged breast-feeding (six months or more) and milk fat content at six months are associated with higher developmental scores at one year of age within a breast-fed population. *Adv Exp Med Biol* 2001;501:137-41.
64. Angelsen NK, Vik T, Jacobsen G, et al. Breast feeding and cognitive development at age 1 and 5 years. *Arch Dis Child* 2001;85(3):183-8.
65. Gomez-Sanchiz M, Canete R, Rodero I, et al. Influence of breast-feeding on mental and psychomotor development. *Clinical Pediatrics* 2003;42(1):35-42.
66. Gomez-Sanchiz M, Canete R, Rodero I, et al. Influence of breast-feeding and parental intelligence on cognitive development in the 24-month-old child. *Clin Pediatr* 2004;43(8):753-61.
67. Mortensen EL, Michaelsen KF, Sanders SA, et al. The association between duration of breastfeeding and adult intelligence.[see comment][erratum appears in *JAMA* 2002 Jun 12;287(22):2946]. *JAMA* 2002;287(18):2365-71.
68. Oddy WH, Kendall GE, Blair E, et al. Breast feeding and cognitive development in childhood: a prospective birth cohort study. *Paediatr Perinat Epidemiol* 2003;17(1):81-90.
69. Oddy WH, Kendall GE, Blair E, et al. Breastfeeding and cognitive development in children. *Adv Exp Med Biol* 2004;554:365-9.
70. Quinn PJ, O'Callaghan M, Williams GM, et al. The effect of breastfeeding on child development at 5 years: a cohort study. *J Paediatr Child Health* 2001;37(5):465-9.
71. Lawlor DA, Najman JM, Batty GD, et al. Early life predictors of childhood intelligence: findings from the Mater-University study of pregnancy

- and its outcomes. *Paediatr Perinat Epidemiol* 2006;20(2):148-62.
72. Evenhouse E, Reilly S. Improved estimates of the benefits of breastfeeding using sibling comparisons to reduce selection bias. *Health Serv Res* 2005;40(6 Pt 1):1781-802.
 73. Agostoni C. Ghrelin, leptin and the neurometabolic axis of breastfed and formula-fed infants. *Acta Paediatr* 2005;94(5):523-5.
 74. Dewey KG. Growth characteristics of breast-fed compared to formula-fed infants. *Biol Neonate* 1998;74(2):94-105.
 75. Demmelmair H, von RJ, Koletzko B. Long-term consequences of early nutrition. [Review] [67 refs]. *Early Hum Dev* 2006;82(8):567-74.
 76. Arenz S, Ruckerl R, Koletzko B, et al. Breast-feeding and childhood obesity--a systematic review. [Review] [51 refs]. *Int J Obes Relat Metab Disord: Journal of the International Association for the Study of Obesity* 2004;28(10):1247-56.
 77. Harder T, Bergmann R, Kallischnigg G, et al. Duration of breastfeeding and risk of overweight: a meta-analysis. *Am J Epidemiol* 2005;162(5):397-403.
 78. Owen CG, Martin RM, Whincup PH, et al. Effect of infant feeding on the risk of obesity across the life course: a quantitative review of published evidence. *Pediatrics* 2005;115(5):1367-77.
 79. Lowe LP, Greenland P, Ruth KJ, et al. Impact of major cardiovascular disease risk factors, particularly in combination, on 22-year mortality in women and men. *Arch Intern Med* 1998;158(18):2007-14.
 80. Stamler J, Daviglus ML, Garside DB, et al. Relationship of baseline serum cholesterol levels in 3 large cohorts of younger men to long-term coronary, cardiovascular, and all-cause mortality and to longevity. *JAMA* 2000;284(3):311-8.
 81. Chobanian AV, Bakris GL, Black HR, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA* 2003;289(19):2560-72.
 82. Owen CG, Whincup PH, Odoki K, et al. Infant feeding and blood cholesterol: a study in adolescents and a systematic review.[see comment]. [Review] [88 refs]. *Pediatrics* 2002;110(3):597-608.
 83. Martin RM, Gunnell D, Smith GD. Breastfeeding in infancy and blood pressure in later life: systematic review and meta-analysis. *Am J Epidemiol* 2005;161(1):15-26.
 84. Owen CG, Whincup PH, Gilg JA, et al. Effect of breast feeding in infancy on blood pressure in later life: systematic review and meta-analysis. [Review] [26 refs]. *BMJ* 2003;327(7425):1189-95.
 85. Martin RM, Davey SG, Mangtani P, et al. Breastfeeding and cardiovascular mortality: the Boyd Orr cohort and a systematic review with meta-analysis. [Review] [25 refs]. *Eur Heart J* 2004;25(9):778-86.
 86. Kleinman R. Towards a "new beginning": dietary fat restrictions in infancy? *Acta Paediatrica* 2000;89(1):2-4.
 87. Virtanen SM, Knip M. Nutritional risk predictors of beta cell autoimmunity and type 1 diabetes at a young age 10046. *Am J Clin Nutr* 2003;78(6):1053-67.
 88. Karjalainen J, Martin JM, Knip M, et al. A bovine albumin peptide as a possible trigger of insulin-dependent diabetes mellitus. *N Engl J Med* 1992;327(5):302-7.
 89. Gerstein HC. Cow's milk exposure and type I diabetes mellitus. A critical overview of the clinical literature. *Diabetes Care* 1994;17(1):13-9.
 90. Norris JM, Scott FW. A meta-analysis of infant diet and insulin-dependent diabetes mellitus: do biases play a role? *Epidemiology* 1996;7(1):87-92.
 91. EURODIAS S. Rapid early growth is associated with increased risk of childhood type 1 diabetes in various European populations. *Diabetes Care* 2002;25(10):1755-60.
 92. Jones ME, Swerdlow AJ, Gill LE, et al. Pre-natal and early life risk factors for childhood onset diabetes mellitus: a record linkage study. *Int J Epidemiol* 1998;27(3):444-9.
 93. McKinney PA, Parslow R, Gurney KA, et al. Perinatal and neonatal determinants of childhood type 1 diabetes. A case-control study in Yorkshire, U.K. *Diabetes Care* 1999;22(6):928-32.
 94. Meloni T, Marinaro AM, Mannazzu MC, et al. IDDM and early infant feeding. Sardinian case-control study. *Diabetes Care* 1997;20(3):340-2.

95. Tai TY, Wang CY, Lin LL, et al. A case-control study on risk factors for Type 1 diabetes in Taipei City. *Diabetes Res Clin Pract* 1998;42(3):197-203.
96. Visalli N, Sebastiani L, Adorisio E, et al. Environmental risk factors for type 1 diabetes in Rome and province. *Arch Dis Child* 2003;88(8):695-8.
97. Lindsay RS, Hanson RL, Bennett PH, et al. Secular trends in birth weight, BMI, and diabetes in the offspring of diabetic mothers. *Diabetes Care* 2000;23(9):1249-54.
98. Taylor JS, Kacmar JE, Nothnagle M, et al. A systematic review of the literature associating breastfeeding with type 2 diabetes and gestational diabetes. *J Am Coll Nutr* 2005;24(5):320-6.
99. Owen CG, Martin RM, Whincup PH, et al. Does breastfeeding influence risk of type 2 diabetes in later life? A quantitative analysis of published evidence. *Am J Clin Nutr* 2006;84(5):1043-54.
100. EPSTEIN MA, ACHONG BG, Barry YM. Vllirus particles in cultured lymphoblastks from Burkitt's lymphoma. *Lancet* 1964;15:702-3.
101. Poesz BJ, Ruscetti FW, Gazdar AF, et al. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci U S A* 1980;77(12):7415-9.
102. Greaves MF. Speculations on the cause of childhood acute lymphoblastic leukemia. *Leukemia* 1988;2(2):120-5.
103. Beral VaUCCSI. Breastfeeding and childhood cancer. *Br J Cancer* 2001;85(11):1685-94.
104. Guise JM, Austin D, Morris CD. Review of case-control studies related to breastfeeding and reduced risk of childhood leukemia. *Pediatrics* 2005;116(5):e724-e731.
105. Davis MK. Review of the evidence for an association between infant feeding and childhood cancer. *Int J Cancer - Suppl* 1998;11:29-33.
106. Kwan ML, Buffler PA, Abrams B, et al. Breastfeeding and the risk of childhood leukemia: a meta-analysis. *Public Health Rep* 2004;119(6):521-35.
107. Shu XO, Linet MS, Steinbuch M, et al. Breast-feeding and risk of childhood acute leukemia. *J Natl Cancer Inst* 1999;91(20):1765-72.
108. Dockerty JD, Skegg DC, Elwood JM, et al. Infections, vaccinations, and the risk of childhood leukaemia. *Br J Cancer* 1999;80(9):1483-9.
109. Rosenbaum PF, Buck GM, Brecher ML. Early child-care and preschool experiences and the risk of childhood acute lymphoblastic leukemia. *Am J Epidemiol* 2000;152(12):1136-44.
110. Eliminate Disparities in Infant Mortality. Office of Minority Health:Centers for Disease Control and Prevention. 2006.
111. Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: a pooled analysis. WHO Collaborative Study Team on the Role of Breastfeeding on the Prevention of Infant Mortality.[see comment][erratum appears in *Lancet* 2000 Mar 25;355(9209):1104]. *Lancet* 2000;355(9202):451-5.
112. Chen A, Rogan WJ. Breastfeeding and the risk of postneonatal death in the United States. *Pediatrics* 2004;113(5):e435-e439.
113. Forste R, Weiss J, Lippincott E. The decision to breastfeed in the United States: does race matter?[see comment]. *Pediatrics* 2001;108(2):291-6.
114. Sudden, Unexplained Infant Death Initiative (SUIDI): Overview. Department for Health and Human Services-Centers for Disease Control and Prevention. 2006.
115. McVea KL, Turner PD, Peppler DK. The role of breastfeeding in sudden infant death syndrome. *J Hum Lact* 2000;16(1):13-20.
116. Findeisen M, Vennemann M, Brinkmann B, et al. German study on sudden infant death (GeSID): design, epidemiological and pathological profile. *Int J Legal Med* 2004;118(3):163-9.
117. Hauck FR, Herman SM, Donovan M, et al. Sleep environment and the risk of sudden infant death syndrome in an urban population: the Chicago Infant Mortality Study. *Pediatrics* 2003;111(5 Part 2):1207-14.
118. Mitchell EA, Tuohy PG, Brunt JM, et al. Risk factors for sudden infant death syndrome following the prevention campaign in New Zealand: a prospective study. *Pediatrics* 1997;100(5):835-40.
119. Schellscheidt J, Ott A, Jorch G. Epidemiological features of sudden infant death after a German

- intervention campaign in 1992. *Eur J Pediatr* 1997;156(8):655-60.
120. Vennemann MM, Findeisen M, Butterfass-Bahloul T, et al. Modifiable risk factors for SIDS in Germany: results of GeSID. *Acta Paediatrica* 2005;94(6):655-60.
 121. Wennergren G, Alm B, Oyen N, et al. The decline in the incidence of SIDS in Scandinavia and its relation to risk-intervention campaigns. Nordic Epidemiological SIDS Study. *Acta Paediatrica* 1997;86(9):963-8.
 122. Brooke H, Gibson A, Tappin D, et al. Case-control study of sudden infant death syndrome in Scotland, 1992-5.[see comment]. *BMJ* 1997;314(7093):1516-20.
 123. Mitchell EA, Taylor BJ, Ford RP, et al. Four modifiable and other major risk factors for cot death: the New Zealand study.[see comment]. *J Paediatr Child Health* 1992;28:Suppl-8.
 124. Fleming PJ, Blair PS, Bacon C, et al. Environment of infants during sleep and risk of the sudden infant death syndrome: results of 1993-5 case-control study for confidential inquiry into stillbirths and deaths in infancy. Confidential Enquiry into Stillbirths and Deaths Regional Coordinators and Researchers.[see comment]. *BMJ* 1996;313(7051):191-5.
 125. Llanos AR, Moss ME, Pinzon MC, et al. Epidemiology of neonatal necrotising enterocolitis: a population-based study. *Paediatr Perinat Epidemiol* 2002;16(4):342-9.
 126. McGuire W, Anthony MY. Formula milk versus term human milk for feeding preterm or low birth weight infants. *Cochrane Database Syst Rev* 2001;(4):CD002971.
 127. Schanler RJ, Lau C, Hurst NM, et al. Randomized trial of donor human milk versus preterm formula as substitutes for mothers' own milk in the feeding of extremely premature infants. *Pediatrics* 2005;116(2):400-6.
 128. Furman L, Taylor G, Minich N, et al. The effect of maternal milk on neonatal morbidity of very low-birth-weight infants. *Arch Pediatr Adolesc Med* 2003;157(1):66-71.
 129. Ronnestad A, Abrahamsen TG, Medbo S, et al. Late-onset septicemia in a Norwegian national cohort of extremely premature infants receiving very early full human milk feeding. *Pediatrics* 2005;115(3):e269-e276.
 130. Lucas A, Gore SM, Cole TJ, et al. Multicentre trial on feeding low birthweight infants: effects of diet on early growth. *Arch Dis Child* 1984;59(8):722-30.
 131. Lucas A, Cole TJ. Breast milk and neonatal necrotising enterocolitis. *Lancet* 1990;336(8730):1519-23.
 132. Dezoete JA, MacArthur BA, Tuck B. Prediction of Bayley and Stanford-Binet scores with a group of very low birthweight children. *Child Care Health Dev* 2003;29(5):367-72.
 133. Hack M, Taylor HG, Drotar D, et al. Poor predictive validity of the Bayley Scales of Infant Development for cognitive function of extremely low birth weight children at school age. *Pediatrics* 2005;116(2):333-41.
 134. Eidelman AI, Feldman R. Positive effect of human milk on neurobehavioral and cognitive development of premature infants. *Adv Exp Med Biol* 2004;554:359-64.
 135. Elgen I, Sommerfelt K, Ellertsen B. Cognitive performance in a low birth weight cohort at 5 and 11 years of age. *Pediatr Neurol* 2003;29(2):111-6.
 136. Feldman R, Eidelman AI. Direct and indirect effects of breast milk on the neurobehavioral and cognitive development of premature infants. *Dev Psychobiol* 2003;43(2):109-19.
 137. Pinelli J, Saigal S, Atkinson SA. Effect of breastmilk consumption on neurodevelopmental outcomes at 6 and 12 months of age in VLBW infants. *Adv Neonat Care* 2003;3(2):76-87.
 138. Vohr BR, Poindexter BB, Dusick AM, et al. Beneficial effects of breast milk in the neonatal intensive care unit on the developmental outcome of extremely low birth weight infants at 18 months of age. *Pediatrics* 2006;118(1):e115-e123.
 139. Bier JA, Oliver T, Ferguson AE, et al. Human milk improves cognitive and motor development of premature infants during infancy. *J Hum Lact* 2002;18(4):361-7.
 140. Smith MM, Durkin M, Hinton VJ, et al. Influence of breastfeeding on cognitive outcomes at age 6-8 years: follow-up of very low birth weight infants. *Am J Epidemiol* 2003;158(11):1075-82.
 141. Horwood LJ, Darlow BA, Mogridge N. Breast milk feeding and cognitive ability at 7-8 years. *Arch Dis Child Fetal Neonatal Ed* 2001;84(1):F23-F27.

142. O'Connor DL, Jacobs J, Hall R, et al. Growth and development of premature infants fed predominantly human milk, predominantly premature infant formula, or a combination of human milk and premature formula. *J Pediatr Gastroenterol Nutr* 2003;37(4):437-46.
143. Smith DE, Lewis CE, Caveny JL, et al. Longitudinal changes in adiposity associated with pregnancy. The CARDIA Study. Coronary Artery Risk Development in Young Adults Study.[see comment]. *JAMA* 1994;271(22):1747-51.
144. Ohlin A, Rossner S. Maternal body weight development after pregnancy. *Int J Obes* 1990;14(2):159-73.
145. Olson CM, Strawderman MS, Hinton PS, et al. Gestational weight gain and postpartum behaviors associated with weight change from early pregnancy to 1 y postpartum. *Int J Obes Relat Metab Disord* 2003;27(1):117-27.
146. Schauburger CW, Rooney BL, Brimer LM. Factors that influence weight loss in the puerperium. *Obstet Gynecol* 1992;79(3):424-9.
147. Greene GW, Smiciklas-Wright H, Scholl TO, et al. Postpartum weight change: how much of the weight gained in pregnancy will be lost after delivery? *Obstet Gynecol* 1988;71(5):701-7.
148. Keppel KG, Taffel SM. Pregnancy-related weight gain and retention: implications of the 1990 Institute of Medicine guidelines.[see comment]. *Am J Public Health* 1993;83(8):1100-3.
149. Fraser AB, Grimes DA. Effect of lactation on maternal body weight: a systematic review. [Review] [57 refs]. *Obstet Gynecol Survey* 2003;58(4):265-9.
150. Butte NF, Hopkinson JM. Body composition changes during lactation are highly variable among women. [Review] [45 refs]. *J Nutr* 1998;128(2:Suppl):Suppl-385S.
151. Dewey KG. Impact of breastfeeding on maternal nutritional status. [Review] [26 refs]. *Adv Exp Med Biol* 2004;554:91-100.
152. Sichieri R, Field AE, Rich-Edwards J, et al. Prospective assessment of exclusive breastfeeding in relation to weight change in women. *Int J Obes Relat Metab Disord* 2003;27(7):815-20.
153. Janney CA, Zhang D, Sowers M. Lactation and weight retention. *Am J Clin Nutr* 1997;66(5):1116-24.
154. Linne Y, Dye L, Barkeling B, et al. Weight development over time in parous women--the SPAWN study--15 years follow-up. *Int J Obes Relat Metab Disord* 2003;27(12):1516-22.
155. Walker L, Freeland-Graves JH, Milani T, et al. Weight and behavioral and psychosocial factors among ethnically diverse, low-income women after childbirth: II. Trends and correlates. *Women Health* 2004;40(2):19-34.
156. Brewer MM, Bates MR, Vannoy LP. Postpartum changes in maternal weight and body fat depots in lactating vs nonlactating women. *Am J Clin Nutr* 1989;49(2):259-65.
157. Ohlin A, Rossner S. Factors related to body weight changes during and after pregnancy: the Stockholm Pregnancy and Weight Development Study. *Obes Res* 1996;4(3):271-6.
158. Haiek LN, Kramer MS, Ciampi A, et al. Postpartum weight loss and infant feeding. *J Am Board Fam Pract* 2001;14(2):85-94.
159. Kjos SL, Henry O, Lee RM, et al. The effect of lactation on glucose and lipid metabolism in women with recent gestational diabetes. *Obstet Gynecol* 1993;82(3):451-5.
160. McManus RM, Cunningham I, Watson A, et al. Beta-cell function and visceral fat in lactating women with a history of gestational diabetes. *Metabolism* 2001;50(6):715-9.
161. Kjos SL, Peters RK, Xiang A, et al. Contraception and the risk of type 2 diabetes mellitus in Latina women with prior gestational diabetes mellitus. *JAMA* 1998;280(6):533-8.
162. Stuebe AM, Rich-Edwards JW, Willett WC, et al. Duration of lactation and incidence of type 2 diabetes. *JAMA* 2005;294(20):2601-10.
163. Sowers M. Pregnancy and lactation as risk factors for subsequent bone loss and osteoporosis. *J Bone Miner Res* 1996;11(8):1052-60.
164. Kovacs CS, Kronenberg HM. Maternal-Fetal Calcium and Bone Metabolism During Pregnancy, Puerperium, and Lactation. *Endocr Rev* 1997;18(6):832-72.
165. Affinito P, Tommaselli GA, di Carlo C, et al. Changes in bone mineral density and calcium metabolism in breastfeeding women: a one year

- follow-up study. *J Clin Endocrinol Metab* 1996;81(6):2314-8.
166. Matsumoto I, Kosha S, Noguchi S, et al. Changes of bone mineral density in pregnant and postpartum women. *J Obstet Gynaecol* 1995;21(5):419-25.
 167. Pearson D, Kaur M, San P, et al. Recovery of pregnancy mediated bone loss during lactation. *Bone* 2004;34(3):570-8.
 168. Sowers M, Corton G, Shapiro B, et al. Changes in bone density with lactation.[see comment]. *JAMA* 1993;269(24):3130.
 169. Lopez JM, Gonzalez G, Reyes V, et al. Bone turnover and density in healthy women during breastfeeding and after weaning. *Osteoporos Int* 1996;6(2):153-9.
 170. Alderman BW, Weiss NS, Daling JR, et al. Reproductive history and postmenopausal risk of hip and forearm fracture. *Am J Epidemiol* 1986;124(2):262-7.
 171. Chan HH, Lau EM, Woo J, et al. Dietary calcium intake, physical activity and the risk of vertebral fracture in Chinese. *Osteoporos Int* 1996;6(3):228-32.
 172. Cumming RG, Klineberg RJ. Breastfeeding and other reproductive factors and the risk of hip fractures in elderly women.[erratum appears in *Int J Epidemiol* 1993 Oct;22(5):962]. *Int J Epidemiol* 1993;22(4):684-91.
 173. Hoffman S, Grisso JA, Kelsey JL, et al. Parity, lactation and hip fracture. *Osteoporos Int* 1993;3(4):171-6.
 174. Kreiger N, Kelsey JL, Holford TR, et al. An epidemiologic study of hip fracture in postmenopausal women. *Am J Epidemiol* 1982;116(1):141-8.
 175. Michaelsson K, Baron JA, Farahmand BY, et al. Influence of parity and lactation on hip fracture risk. *Am J Epidemiol* 2001;153(12):1166-72.
 176. Hansen MA, Overgaard K, Riis BJ, et al. Potential risk factors for development of postmenopausal osteoporosis--examined over a 12-year period. *Osteoporos Int* 1991;1(2):95-102.
 177. Matsushita H, Kurabayashi T, Tomita M, et al. The effect of multiple pregnancies on lumbar bone mineral density in Japanese women. *Calcif Tissue Int* 2002;71(1):10-3.
 178. Sowers MR, Clark MK, Hollis B, et al. Radial bone mineral density in pre- and perimenopausal women: a prospective study of rates and risk factors for loss. *J Bone Miner Res* 1992;7(6):647-57.
 179. Uusi-Rasi K, Sievanen H, Pasanen M, et al. Association of physical activity and calcium intake with the maintenance of bone mass in premenopausal women. *Osteoporos Int* 2002;13(3):211-7.
 180. O'Hara MW, Swain AM. Rates and risk of postpartum depression - a meta-analysis. *Int Rev Psychiatry* 1996;8(1):37-54.
 181. Warner R, Appleby L, Whitton A, et al. Demographic and obstetric risk factors for postnatal psychiatric morbidity. *Br J Psychiatry* 1996;168(5):607-11.
 182. Chaudron LH, Klein MH, Remington P, et al. Predictors, prodromes and incidence of postpartum depression. *J Psychosom Obstet Gynecol* 2001;22(2):103-12.
 183. Cooper PJ, Murray L, Stein A. Psychosocial factors associated with the early termination of breast-feeding. *J Psychosom Res* 1993;37(2):171-6.
 184. Hannah P, Adams D, Lee A, et al. Links between early post-partum mood and post-natal depression. *Br J Psychiatry* 1992;160:777-80.
 185. Henderson JJ, Evans SF, Straton JA, et al. Impact of postnatal depression on breastfeeding duration.[erratum appears in *Birth*. 2004 Mar;31(1):76]. *Birth* 2003;30(3):175-80.
 186. Seimyr L, Edhborg M, Lundh W, et al. In the shadow of maternal depressed mood: experiences of parenthood during the first year after childbirth. *J Psychosom Obstet Gynecol* 2004;25(1):23-34.
 187. American Cancer Society: Cancer Facts and Figures 2006. Atlanta, GA. American Cancer Society, 2006.
 188. Bernier MO, PluBureau G, Bossard N, et al. Breastfeeding and risk of breast cancer: a metaanalysis of published studies. *Hum Reprod Update* 2000;6(4):374-86.
 189. Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and

- 96973 women without the disease.[see comment]. [Review] [68 refs]. *Lancet* 2002;360(9328):187-95.
190. Lipworth L, Bailey LR, Trichopoulos D. History of breast-feeding in relation to breast cancer risk: a review of the epidemiologic literature.[see comment]. [Review] [43 refs]. *J Natl Cancer Inst* 2000;92(4):302-12.
 191. Gammon MD, Neugut AI, Santella RM, et al. The Long Island Breast Cancer Study Project: description of a multi-institutional collaboration to identify environmental risk factors for breast cancer. *Breast Cancer Res Treat* 2002;74(3):235-54.
 192. Jernstrom H, Lubinski J, Lynch HT, et al. Breast-feeding and the risk of breast cancer in BRCA1 and BRCA2 mutation carriers. *J Natl Cancer Inst* 2004;96(14):1094-8.
 193. Lee SY, Kim MT, Kim SW, et al. Effect of lifetime lactation on breast cancer risk: a Korean women's cohort study. *Int J Cancer* 2003;105(3):390-3.
 194. Riman T, Dickman PW, Nilsson S, et al. Risk factors for invasive epithelial ovarian cancer: results from a Swedish case-control study. *Am J Epidemiol* 2002;156(4):363-73.
 195. Chiaffarino F, Pelucchi C, Negri E, et al. Breastfeeding and the risk of epithelial ovarian cancer in an Italian population. *Gyneco Oncol* 2005;98(2):304-8.
 196. Cramer DW, Hutchison GB, Welch WR, et al. Determinants of ovarian cancer risk. I. Reproductive experiences and family history. *J Natl Cancer Inst* 1983;71(4):711-6.
 197. Gregg S, Parazzini F, Paratore MP, et al. Risk factors for ovarian cancer in central Italy. *Gynecol Oncol* 2000;79(1):50-4.
 198. Gwinn ML, Lee NC, Rhodes PH, et al. Pregnancy, breast feeding, and oral contraceptives and the risk of epithelial ovarian cancer. *J Clin Epidemiol* 1990;43(6):559-68.
 199. Hartge P, Schiffman MH, Hoover R, et al. A case-control study of epithelial ovarian cancer. *Am J Obstet Gynecol* 1989;161(1):10-6.
 200. Modugno F, Ness RB, Wheeler JE. Reproductive risk factors for epithelial ovarian cancer according to histologic type and invasiveness. *Ann Epidemiol* 2001 Nov; 11(8):568-74.
 201. Ness RB, Grisso JA, Cottreau C, et al. Factors related to inflammation of the ovarian epithelium and risk of ovarian cancer.[see comment]. *Epidemiology* 2000;11(2):111-7.
 202. Risch HA, Weiss NS, Lyon JL, et al. Events of reproductive life and the incidence of epithelial ovarian cancer. *Am J Epidemiol* 1983;117(2):128-39.
 203. Risch HA, Marrett LD, Howe GR. Parity, contraception, infertility, and the risk of epithelial ovarian cancer. *Am J Epidemiol* 1994;140(7):585-97.
 204. Siskind V, Green A, Bain C, et al. Breastfeeding, menopause, and epithelial ovarian cancer. *Epidemiology* 1997;8(2):188-91.
 205. Titus-Ernstoff L, Perez K, Cramer DW, et al. Menstrual and reproductive factors in relation to ovarian cancer risk. *Br J Cancer* 2001;84(5):714-21.
 206. Tung KH, Goodman MT, Wu AH, et al. Reproductive factors and epithelial ovarian cancer risk by histologic type: a multiethnic case-control study. *Am J Epidemiol* 2003;158(7):629-38.
 207. Tung KH, Wilkens LR, Wu AH, et al. Effect of anovulation factors on pre- and postmenopausal ovarian cancer risk: revisiting the incessant ovulation hypothesis. *Am J Epidemiol* 2005;161(4):321-9.
 208. West RO. Epidemiologic study of malignancies of the ovaries. *Cancer* 1966;19(7):1001-7.
 209. Wynder EL, Dodo H, Barber HR. Epidemiology of cancer of the ovary. *Cancer* 1969;23(2):352-70.
 210. Yen ML, Yen BL, Bai CH, et al. Risk factors for ovarian cancer in Taiwan: a case-control study in a low-incidence population. *Gynecologic Oncology* 2003;89(2):318-24.
 211. John EM, Whittemore AS, Harris R, et al. Characteristics relating to ovarian cancer risk: collaborative analysis of seven U.S. case-control studies. Epithelial ovarian cancer in black women. Collaborative Ovarian Cancer Group. *Journal of the National Cancer Institute* 1993;85(2):142-7.
 212. Whittemore AS, Harris R, Itnyre J. Characteristics relating to ovarian cancer risk: collaborative analysis of 12 US case-control studies. II. Invasive epithelial ovarian cancers in

white women. Collaborative Ovarian Cancer Group.[see comment]. *Am J Epidemiol* 1992;136(10):1184-203.

213. Modugno F, Moslehi R, Ness RB, et al. Reproductive factors and ovarian cancer risk in Jewish BRCA1 and BRCA2 mutation carriers (United States). *Cancer Causes Control* 2003;14(5):439-46.
214. Kramer MS, Kakuma R. The optimal duration of exclusive breastfeeding: a systematic review. *Adv Exp Med Biol* 2004;554:63-77.
215. Bland RM, Rollins NC, Solarsh G, et al. Maternal recall of exclusive breast feeding duration. *Arch Dis Child* 2003;88(9):778-83.
216. Neville MC, Keller RP, Seacat J, et al. Studies on human lactation. I. Within-feed and between-breast variation in selected components of human milk. *Am J Clin Nutr* 1984;40(3):635-46.
217. Nommsen LA, Lovelady CA, Heinig MJ, et al. Determinants of energy, protein, lipid, and lactose concentrations in human milk during the first 12 mo of lactation: the DARLING Study. *Am J Clin Nutr* 1991;53(2):457-65.
218. Newburg DS, Ruiz-Palacios GM, Morrow AL. Human milk glycans protect infants against enteric pathogens. *Annu Rev Nutr* 2005;25:37-58.

Bibliography

Websites referenced in report:

Asthma prevalence in: Health, United States, 2005 With Chartbook on Trends in the Health of Americans. Centers for Disease Control, National Center for Health Statistics Web site. Available at: <http://www.cdc.gov/nchs/data/hs/hs05.pdf>. Last accessed on March 6, 2007.

Diabetes in Youth. National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention. Web site. Available at: <http://www.cdc.gov/diabetes/pubs/factsheets/search.htm>. Last accessed on March 6, 2006.

Healthy People 2010 – Breastfeeding Practices. Results from the 2005 National Immunization Survey. Available at: http://www.cdc.gov/breastfeeding/data/NIS_data/data_2005.htm. Last accessed on March 6, 2007.

HHS Blueprint for Action on Breastfeeding.U.S. Department of Health and Human Services, Office of Women's Health Womenshealth.gov.Web site. Available at: <http://www.womenshealth.gov/Breastfeeding/bluprntbk2.pdf>. Last accessed on March 7, 2007

Mortality Data. National Center for Health Statistics Web site: Available at: <http://www.cdc.gov/nchs/deaths.htm>. Last accessed March 6, 2007.

National Immunization Survey, 2005. Centers for Disease Control and Prevention Web site. Available at: http://www.cdc.gov/breastfeeding/data/NIS_data/data_2005.htm. Last accessed on March 6, 2007.

Prevalence of Overweight among Children and Adolescents: United States, 1999-2002. National Center of Health Statistics, Centers for Disease Control and Prevention Web site. Available at: <http://www.cdc.gov/nchs/products/pubs/pubd/hestats/overwght99.htm>. Last accessed on March 6, 2007.

WIC Program Guide to WIC Women, Infants, and Children Program. Available at: <http://www.wicprogram.org/>. Last accessed on March 6, 2007.

World Bank Data & Statistics Page. World Bank Web site. Available at: http://web.worldbank.org/WBSITE/EXTERNAL/DATASTATISTICS/0,,contentMDK:20421402~pagePK:64133150~piPK:64133175~theSitePK:239419,00.html#High_income. Last accessed on March 6, 2007.

References:

Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: a pooled analysis. WHO Collaborative Study Team on the Role of Breastfeeding on the Prevention of Infant Mortality.[see comment][erratum appears in Lancet 2000 Mar 25;355(9209):1104]. Lancet 2000 Feb 5;355(9202):451-5.

American Cancer Society. Cancer Facts and Figures 2006. 2006a. Atlanta, GA, American Cancer Society, 2006.

Eliminate Disparities in Infant Mortality. 2006d. Office of Minority Health:Centers for Disease Control and Prevention.

Sudden, Unexplained Infant Death Initiative (SUIDI): Overview. 2006e. Department for Health and Human Services-Centers for Disease Control and Prevention.

Affinito P, Tommaselli GA, di Carlo C, et al. Changes in bone mineral density and calcium metabolism in breastfeeding women: a one year follow-up study. J Clin Endocrinol Metab 1996 Jun;81(6):2314-8.

Agostoni C. Ghrelin, leptin and the neurometabolic axis of breastfed and formula-fed infants. Acta Paediatr 2005 May;94(5):523-5.

Agostoni C, Marangoni F, Giovannini M, et al. Prolonged breast-feeding (six months or more) and milk fat content at six months are associated with higher developmental scores at one year of age within a breast-fed population. Adv Exp Med Biol 2001;501:137-41.

Alderman BW, Weiss NS, Daling JR, et al. Reproductive history and postmenopausal risk of hip and forearm fracture. Am J Epidemiol 1986 Aug;124(2):262-7.

Alho OP, Laara, Oja H. How should relative risk estimates for acute otitis media in children aged less than 2 years be perceived? J Clin Epidemiol 1996 Jan;49(1):9-14.

Alho OP, Laara E, Oja H. Public health impact of various risk factors for acute otitis media in northern Finland. Am J Epidemiol 1996 Jun 1;143(11):1149-56.

Altman DG, Schulz KF, Moher D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med 2001 Apr 17;134(8):663-94.

American Academy of Pediatrics Committee on Nutrition. Pediatric Nutrition Handbook. Kleinman, R. E. Fifth. 2004.

- Anderson JW, Johnstone BM, Remley DT. Breast-feeding and cognitive development: a meta-analysis. *Am J Clin Nutr* 1999 Oct;70(4):525-35.
- Angelsen NK, Vik T, Jacobsen G, et al. Breast feeding and cognitive development at age 1 and 5 years. *Arch Dis Child* 2001 Sep;85(3):183-8.
- Arenz S, Ruckerl R, Koletzko B, et al. Breast-feeding and childhood obesity--a systematic review. [Review] [51 refs]. *Int J Obes Relat Metab Disord: Journal of the International Association for the Study of Obesity* 2004 Oct;28(10):1247-56.
- Auinger P, Lanphear BP, Kalkwarf HJ, et al. Trends in otitis media among children in the United States. *Pediatrics* 2003 Sep;112(3 Pt 1):514-20.
- Bachrach VR, Schwarz E, Bachrach LR. Breastfeeding and the risk of hospitalization for respiratory disease in infancy: a meta-analysis. *Arch Pediatr Adolesc Med* 2003 Mar;157(3):237-43.
- Bauchner H, Leventhal JM, Shapiro ED. Studies of breast-feeding and infections. How good is the evidence? *JAMA* 1986 Aug 15;256(7):887-92.
- Beral VaUCCSI. Breastfeeding and childhood cancer. *Br J Cancer* 2001 Nov 30;85(11):1685-94.
- Bernier MO, PluBureau G, Bossard N, et al. Breastfeeding and risk of breast cancer: a metaanalysis of published studies. *Hum Reprod Update* 2000 Jul;6(4):374-86.
- Bier JA, Oliver T, Ferguson AE, et al. Human milk improves cognitive and motor development of premature infants during infancy. *J Hum Lact* 2002 Nov;18(4):361-7.
- Bland RM, Rollins NC, Solarsh G, et al. Maternal recall of exclusive breast feeding duration. *Arch Dis Child* 2003 Sep;88(9):778-83.
- Bluestone, CD and Klein, JO. Otitis media in infants and children. 2nd edition. 1995. WB Saunders.
- Braun-Fahrlander C, Riedler J, Herz U, et al. Environmental exposure to endotoxin and its relation to asthma in school-age children. *N Engl J Med* 2002 Sep 19;347(12):869-77.
- Brewer MM, Bates MR, Vannoy LP. Postpartum changes in maternal weight and body fat depots in lactating vs nonlactating women. *Am J Clin Nutr* 1989 Feb;49(2):259-65.
- Brooke H, Gibson A, Tappin D, et al. Case-control study of sudden infant death syndrome in Scotland, 1992-5.[see comment]. *BMJ* 1997 May 24;314(7093):1516-20.
- Burgess SW, Dakin CJ, O'Callaghan MJ. Breastfeeding does not increase the risk of asthma at 14 years. *Pediatrics* 2006 Apr;117(4):e787-e792.
- Butte NF, Hopkinson JM. Body composition changes during lactation are highly variable among women. [Review] [45 refs]. *J Nutr* 1998 Feb;128(2:Suppl):Suppl-385S.
- Chan HH, Lau EM, Woo J, et al. Dietary calcium intake, physical activity and the risk of vertebral fracture in Chinese. *Osteoporos Int* 1996;6(3):228-32.
- Chaudron LH, Klein MH, Remington P, et al. Predictors, prodromes and incidence of postpartum depression. *J Psychosom Obstet Gynecol* 2001 Jun;22(2):103-12.
- Chen A, Rogan WJ. Breastfeeding and the risk of postneonatal death in the United States. *Pediatrics* 2004 May;113(5):e435-e439.
- Chiapparino F, Pelucchi C, Negri E, et al. Breastfeeding and the risk of epithelial ovarian cancer in an Italian population. *Gynecol Oncol* 2005 Aug;98(2):304-8.
- Chien PF, Howie PW. Breast milk and the risk of opportunistic infection in infancy in industrialized and non-industrialized settings. [Review] [75 refs]. *Adv Nutr Res* 2001;10:69-104.
- Chobanian AV, Bakris GL, Black HR, et al. The Seventh Report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: the JNC 7 report. *JAMA* 2003;289(19):2560-72.
- Chua S, Arulkumaran S, Lim I, et al. Influence of breastfeeding and nipple stimulation on postpartum uterine activity. *Br J Obstet Gynaecol* 1994 Sep;101(9):804-5.
- Collaborative Group on Hormonal Factors in Breast Cancer. Breast cancer and breastfeeding: collaborative reanalysis of individual data from 47 epidemiological studies in 30 countries, including 50302 women with breast cancer and 96973 women without the disease.[see comment][comment]. [Review] [68 refs]. *Lancet* 2002 Jul 20;360(9328):187-95.
- Cooper PJ, Murray L, Stein A. Psychosocial factors associated with the early termination of breast-feeding. *J Psychosom Res* 1993;37(2):171-6.
- Cramer DW, Hutchison GB, Welch WR, et al. Determinants of ovarian cancer risk. I. Reproductive experiences and family history. *J Natl Cancer Inst* 1983 Oct;71(4):711-6.
- Cumming RG, Klineberg RJ. Breastfeeding and other reproductive factors and the risk of hip fractures in elderly women.[erratum appears in *Int J Epidemiol* 1993 Oct;22(5):962]. *Int J Epidemiol* 1993 Aug;22(4):684-91.

- Davis MK. Review of the evidence for an association between infant feeding and childhood cancer. *Int J Cancer - Suppl* 1998;11:29-33.
- Demmelmair H, von RJ, Koletzko B. Long-term consequences of early nutrition. [Review] [67 refs]. *Early Hum Dev* 2006 Aug;82(8):567-74.
- Der G, Batty D, Deary IJ. Effect of breast feeding on intelligence in children: prospective study, sibling pairs analysis, and meta-analysis. *BMJ* 2006 Oct 4c.
- DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986 Sep;7(3):177-88.
- Dewey KG. Growth characteristics of breast-fed compared to formula-fed infants. *Biol Neonate* 1998;74(2):94-105.
- Dewey KG. Impact of breastfeeding on maternal nutritional status. [Review] [26 refs]. *Adv Exp Med Biol* 2004;554:91-100.
- Dewey KG, Heinig MJ, Nommsen LA. Maternal weight-loss patterns during prolonged lactation. *Am J Clin Nutr* 1993 Aug;58(2):162-6.
- Dezoete JA, MacArthur BA, Tuck B. Prediction of Bayley and Stanford-Binet scores with a group of very low birthweight children. *Child Care Health Dev* 2003 Sep;29(5):367-72.
- Dockerty JD, Skegg DC, Elwood JM, et al. Infections, vaccinations, and the risk of childhood leukaemia. *Br J Cancer* 1999 Jul;80(9):1483-9.
- Drane DL, Logemann JA. A critical evaluation of the evidence on the association between type of infant feeding and cognitive development. *Paediatr Perinat Epidemiol* 2000 Oct;14(4):349-56.
- Duffy LC, Faden H, Wasielewski R, et al. Exclusive breastfeeding protects against bacterial colonization and day care exposure to otitis media. *Pediatrics* 1997 Oct;100(4):E7.
- Duncan B, Ey J, Holberg CJ, et al. Exclusive breast-feeding for at least 4 months protects against otitis media.[see comment]. *Pediatrics* 1993 May;91(5):867-72.
- Eder W, Ege MJ, von ME. The asthma epidemic. *N Engl J Med* 2006 Nov 23;355(21):2226-35.
- Eidelman AI, Feldman R. Positive effect of human milk on neurobehavioral and cognitive development of premature infants. *Adv Exp Med Biol* 2004;554:359-64.
- Elgen I, Sommerfelt K, Ellertsen B. Cognitive performance in a low birth weight cohort at 5 and 11 years of age. *Pediatr Neurol* 2003 Aug;29(2):111-6.
- EPSTEIN MA, ACHONG BG, Barry YM. Virus particles in cultured lymphoblasts from Burkitt's lymphoma. *Lancet* 1964 Mar 28;15:702-3.
- EURODIAS S. Rapid early growth is associated with increased risk of childhood type 1 diabetes in various European populations. *Diabetes Care* 2002 Oct;25(10):1755-60.
- Evenhouse E, Reilly S. Improved estimates of the benefits of breastfeeding using sibling comparisons to reduce selection bias. *Health Serv Res* 2005 Dec;40(6 Pt 1):1781-802.
- Feachem RG, Koblinsky MA. Interventions for the control of diarrhoeal diseases among young children: promotion of breast-feeding 10045. *Bull World Health Organ* 1984;62(2):271-91.
- Feldman R, Eidelman AI. Direct and indirect effects of breast milk on the neurobehavioral and cognitive development of premature infants. *Dev Psychobiol* 2003 Sep;43(2):109-19.
- Findeisen M, Vennemann M, Brinkmann B, et al. German study on sudden infant death (GeSID): design, epidemiological and pathological profile. *Int J Legal Med* 2004 Jun;118(3):163-9.
- Fleming PJ, Blair PS, Bacon C, et al. Environment of infants during sleep and risk of the sudden infant death syndrome: results of 1993-5 case-control study for confidential inquiry into stillbirths and deaths in infancy. Confidential Enquiry into Stillbirths and Deaths Regional Coordinators and Researchers.[see comment]. *BMJ* 1996 Jul 27;313(7051):191-5.
- Forste R, Weiss J, Lippincott E. The decision to breastfeed in the United States: does race matter?[see comment]. *Pediatrics* 2001 Aug;108(2):291-6.
- Fraser AB, Grimes DA. Effect of lactation on maternal body weight: a systematic review. [Review] [57 refs]. *Obstet Gynecol Survey* 2003 Apr;58(4):265-9.
- Furman L, Taylor G, Minich N, et al. The effect of maternal milk on neonatal morbidity of very low-birth-weight infants. *Arch Pediatr Adolesc Med* 2003 Jan;157(1):66-71.
- Gammon MD, Neugut AI, Santella RM, et al. The Long Island Breast Cancer Study Project: description of a multi-institutional collaboration to identify environmental risk factors for breast cancer. *Breast Cancer Res Treat* 2002 Jun;74(3):235-54.
- Gartner LM, Morton J, Lawrence RA, et al. Breastfeeding and the use of human milk. *Pediatrics* 2005 Feb;115(2):496-506.

- Gdalevich M, Mimouni D, David M, et al. Breast-feeding and the onset of atopic dermatitis in childhood: a systematic review and meta-analysis of prospective studies. *J Am Acad Dermatol* 2001 Oct;45(4):520-7.
- Gdalevich M, Mimouni D, Mimouni M. Breast-feeding and the risk of bronchial asthma in childhood: a systematic review with meta-analysis of prospective studies. *J Pediatr* 2001 Aug;139(2):261-6.
- Gerstein HC. Cow's milk exposure and type I diabetes mellitus. A critical overview of the clinical literature. *Diabetes Care* 1994 Jan;17(1):13-9.
- Gillman MW, Rifas-Shiman SL, Camargo CA, Jr., et al. Risk of overweight among adolescents who were breastfed as infants. *JAMA* 2001 May 16;285(19):2461-7.
- Gomez-Sanchiz M, Canete R, Rodero I, et al. Influence of breast-feeding on mental and psychomotor development. *Clin Pediatr* 2003 Jan;42(1):35-42.
- Gomez-Sanchiz M, Canete R, Rodero I, et al. Influence of breast-feeding and parental intelligence on cognitive development in the 24-month-old child. *Clin Pediatr* 2004 Oct;43(8):753-61.
- Greaves MF. Speculations on the cause of childhood acute lymphoblastic leukemia. *Leukemia* 1988 Feb;2(2):120-5.
- Greene GW, Smiciklas-Wright H, Scholl TO, et al. Postpartum weight change: how much of the weight gained in pregnancy will be lost after delivery? *Obstet Gynecol* 1988 May;71(5):701-7.
- Greggi S, Parazzini F, Paratore MP, et al. Risk factors for ovarian cancer in central Italy. *Gynecol Oncol* 2000 Oct;79(1):50-4.
- Gross SJ. Growth and biochemical response of preterm infants fed human milk or modified infant formula. *N Engl J Med* 1983 Feb 3;308(5):237-41.
- Grummer-Strawn LM, Mei Z. Does breastfeeding protect against pediatric overweight? Analysis of longitudinal data from the Centers for Disease Control and Prevention Pediatric Nutrition Surveillance System. *Pediatrics* 2004 Feb;113(2):e81-e86.
- Guisse JM, Austin D, Morris CD. Review of case-control studies related to breastfeeding and reduced risk of childhood leukemia. *Pediatrics* 2005 Nov;116(5):e724-e731.
- Gwinn ML, Lee NC, Rhodes PH, et al. Pregnancy, breast feeding, and oral contraceptives and the risk of epithelial ovarian cancer. *J Clin Epidemiol* 1990;43(6):559-68.
- Hack M, Taylor HG, Drotar D, et al. Poor predictive validity of the Bayley Scales of Infant Development for cognitive function of extremely low birth weight children at school age. *Pediatrics* 2005 Aug;116(2):333-41.
- Haiek LN, Kramer MS, Ciampi A, et al. Postpartum weight loss and infant feeding. *J Am Board Fam Pract* 2001 Mar;14(2):85-94.
- Hannah P, Adams D, Lee A, et al. Links between early post-partum mood and post-natal depression. *Br J Psychiatry* 1992 Jun;160:777-80.
- Hansen MA, Overgaard K, Riis BJ, et al. Potential risk factors for development of postmenopausal osteoporosis--examined over a 12-year period. *Osteoporos Int* 1991 Feb;1(2):95-102.
- Harder T, Bergmann R, Kallischnigg G, et al. Duration of breastfeeding and risk of overweight: a meta-analysis. *Am J Epidemiol* 2005 Sep 1;162(5):397-403.
- Hartge P, Schiffman MH, Hoover R, et al. A case-control study of epithelial ovarian cancer. *Am J Obstet Gynecol* 1989 Jul;161(1):10-6.
- Hauck FR, Herman SM, Donovan M, et al. Sleep environment and the risk of sudden infant death syndrome in an urban population: the Chicago Infant Mortality Study. *Pediatrics* 2003 May;111(5 Part 2):1207-14.
- Henderson JJ, Evans SF, Straton JA, et al. Impact of postnatal depression on breastfeeding duration.[erratum appears in *Birth*. 2004 Mar;31(1):76]. *Birth* 2003 Sep;30(3):175-80.
- Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. *BMJ* 2003 Sep 6;327(7414):557-60.
- Higgins JP, Whitehead A, Turner RM, et al. Meta-analysis of continuous outcome data from individual patients. *Stat Med* 2001 Aug 15;20(15):2219-41.
- Hoffman S, Grisso JA, Kelsey JL, et al. Parity, lactation and hip fracture. *Osteoporos Int* 1993 Jul;3(4):171-6.
- Horwood LJ, Darlow BA, Mogridge N. Breast milk feeding and cognitive ability at 7-8 years. *Arch Dis Child Fetal Neonat Ed* 2001 Jan;84(1):F23-F27.
- Howie PW, Forsyth JS, Ogston SA, et al. Protective effect of breast feeding against infection. *BMJ* 1990 Jan 6;300(6716):11-6.
- Imhoff B, Morse D, Shiferaw B, et al. Burden of self-reported acute diarrheal illness in FoodNet surveillance areas, 1998-1999. *Clin Infect Dis* 2004 Apr 15;38 Suppl 3:S219-S226.

- Jain A, Concato J, Leventhal JM. How good is the evidence linking breastfeeding and intelligence? *Pediatrics* 2002 Jun;109(6):1044-53.
- Janney CA, Zhang D, Sowers M. Lactation and weight retention. *Am J Clin Nutr* 1997 Nov;66(5):1116-24.
- Jernstrom H, Lubinski J, Lynch HT, et al. Breast-feeding and the risk of breast cancer in BRCA1 and BRCA2 mutation carriers. *J Natl Cancer Inst* 2004 Jul 21;96(14):1094-8.
- John EM, Whittemore AS, Harris R, et al. Characteristics relating to ovarian cancer risk: collaborative analysis of seven U.S. case-control studies. Epithelial ovarian cancer in black women. Collaborative Ovarian Cancer Group. *J Natl Cancer Inst* 1993 Jan;85(2):142-7.
- Johnson DL, Swank PR, Howie VM, et al. Breast feeding and children's intelligence. *Psychol Rep* 1996 Dec;79(3 Part 2):1179-85.
- Jones ME, Swerdlow AJ, Gill LE, et al. Pre-natal and early life risk factors for childhood onset diabetes mellitus: a record linkage study. *Int J Epidemiol* 1998 Jun;27(3):444-9.
- Karjalainen J, Martin JM, Knip M, et al. A bovine albumin peptide as a possible trigger of insulin-dependent diabetes mellitus. *N Engl J Med* 1992 Jul 30;327(5):302-7.
- Keppel KG, Taffel SM. Pregnancy-related weight gain and retention: implications of the 1990 Institute of Medicine guidelines.[see comment]. *Am J Public Health* 1993 Aug;83(8):1100-3.
- Kjos SL, Henry O, Lee RM, et al. The effect of lactation on glucose and lipid metabolism in women with recent gestational diabetes. *Obstet Gynecol* 1993 Sep;82(3):451-5.
- Kjos SL, Peters RK, Xiang A, et al. Contraception and the risk of type 2 diabetes mellitus in Latina women with prior gestational diabetes mellitus. *JAMA* 1998 Aug 12;280(6):533-8.
- Kleinman R. Towards a "new beginning": dietary fat restrictions in infancy? *Acta Paediatrica* 2000 Jan;89(1):2-4.
- Kovacs CS, Kronenberg HM. Maternal-Fetal Calcium and Bone Metabolism During Pregnancy, Puerperium, and Lactation. *Endocr Rev* 1997 Dec 1;18(6):832-72.
- Kramer MS. Does breast feeding help protect against atopic disease? Biology, methodology, and a golden jubilee of controversy. *J Pediatr* 1988 Feb;112(2):181-90.
- Kramer MS, Chalmers B, Hodnett ED, et al. Promotion of Breastfeeding Intervention Trial (PROBIT): a randomized trial in the Republic of Belarus. *JAMA* 2001 Jan 24;285(4):413-20.
- Kramer MS, Kakuma R. The optimal duration of exclusive breastfeeding: a systematic review. *Adv Exp Med Biol* 2004;554:63-77.
- Kreiger N, Kelsey JL, Holford TR, et al. An epidemiologic study of hip fracture in postmenopausal women. *Am J Epidemiol* 1982 Jul;116(1):141-8.
- Kull I, Almqvist C, Lilja G, et al. Breast-feeding reduces the risk of asthma during the first 4 years of life.[see comment]. *J Allergy Clin Immunol* 114(4):755-60, 2004 Oct;114(4):755-60.
- Kwan ML, Buffler PA, Abrams B, et al. Breastfeeding and the risk of childhood leukemia: a meta-analysis. *Public Health Rep* 2004 Nov;119(6):521-35.
- Lawlor DA, Najman JM, Batty GD, et al. Early life predictors of childhood intelligence: findings from the Mater-University study of pregnancy and its outcomes. *Paediatr Perinat Epidemiol* 2006 Mar;20(2):148-62.
- Lee SY, Kim MT, Kim SW, et al. Effect of lifetime lactation on breast cancer risk: a Korean women's cohort study. *Int J Cancer* 2003 Jun 20;105(3):390-3.
- Li R, Darling N, Maurice E, et al. Breastfeeding rates in the United States by characteristics of the child, mother, or family: the 2002 National Immunization Survey. *Pediatrics* 2005 Jan;115(1):e31-e37.
- Lindsay RS, Hanson RL, Bennett PH, et al. Secular trends in birth weight, BMI, and diabetes in the offspring of diabetic mothers. *Diabetes Care* 2000 Sep 1;23(9):1249-54.
- Linne Y, Dye L, Barkeling B, et al. Weight development over time in parous women--the SPAWN study--15 years follow-up. *Int J Obes Relat Metab Disord* 2003 Dec;27(12):1516-22.
- Lipworth L, Bailey LR, Trichopoulos D. History of breastfeeding in relation to breast cancer risk: a review of the epidemiologic literature.[see comment]. [Review] [43 refs]. *J Natl Cancer Inst* 2000 Feb 16;92(4):302-12.
- Llanos AR, Moss ME, Pinzon MC, et al. Epidemiology of neonatal necrotizing enterocolitis: a population-based study. *Paediatr Perinat Epidemiol* 2002 Oct;16(4):342-9.
- Lopez JM, Gonzalez G, Reyes V, et al. Bone turnover and density in healthy women during breastfeeding and after weaning. *Osteoporos Int* 1996;6(2):153-9.
- Lowe LP, Greenland P, Ruth KJ, et al. Impact of major cardiovascular disease risk factors, particularly in

- combination, on 22-year mortality in women and men. *Arch Intern Med* 1998 Oct 12;158(18):2007-14.
- Lucas A, Cole TJ. Breast milk and neonatal necrotising enterocolitis. *Lancet* 1990 Dec 22;336(8730):1519-23.
- Lucas A, Gore SM, Cole TJ, et al. Multicentre trial on feeding low birthweight infants: effects of diet on early growth. *Arch Dis Child* 1984 Aug;59(8):722-30.
- Martin RM, Davey SG, Mangtani P, et al. Breastfeeding and cardiovascular mortality: the Boyd Orr cohort and a systematic review with meta-analysis. [Review] [25 refs]. *Eur Heart J* 2004 May;25(9):778-86.
- Martin RM, Gunnell D, Smith GD. Breastfeeding in infancy and blood pressure in later life: systematic review and meta-analysis. *Am J Epidemiol* 2005 Jan 1;161(1):15-26.
- Matsumoto I, Kosha S, Noguchi S, et al. Changes of bone mineral density in pregnant and postpartum women. *J Obstet Gynaecol* 1995 Oct;21(5):419-25.
- Matsushita H, Kurabayashi T, Tomita M, et al. The effect of multiple pregnancies on lumbar bone mineral density in Japanese women. *Calcif Tissue Int* 2002 Jul;71(1):10-3.
- McGuire W, Anthony MY. Formula milk versus term human milk for feeding preterm or low birth weight infants. *Cochrane Database Syst Rev* 2001;(4):CD002971.
- McKinney PA, Parslow R, Gurney KA, et al. Perinatal and neonatal determinants of childhood type 1 diabetes. A case-control study in Yorkshire, U.K. *Diabetes Care* 1999 Jun;22(6):928-32.
- McManus RM, Cunningham I, Watson A, et al. Beta-cell function and visceral fat in lactating women with a history of gestational diabetes. *Metabolism* 2001 Jun;50(6):715-9.
- McVea KL, Turner PD, Peppler DK. The role of breastfeeding in sudden infant death syndrome. *J Hum Lact* 2000 Feb;16(1):13-20.
- Meloni T, Marinaro AM, Mannazzu MC, et al. IDDM and early infant feeding. Sardinian case-control study. *Diabetes Care* 1997;(3):340-2.
- Michaelsson K, Baron JA, Farahmand BY, et al. Influence of parity and lactation on hip fracture risk. *Am J Epidemiol* 2001 Jun 15;153(12):1166-72.
- Mitchell EA, Taylor BJ, Ford RP, et al. Four modifiable and other major risk factors for cot death: the New Zealand study.[see comment]. *J Paediatr Child Health* 1992;28:Suppl-8.
- Mitchell EA, Tuohy PG, Brunt JM, et al. Risk factors for sudden infant death syndrome following the prevention campaign in New Zealand: a prospective study. *Pediatrics* 1997 Nov;100(5):835-40.
- Modugno F, Moslehi R, Ness RB, et al. Reproductive factors and ovarian cancer risk in Jewish BRCA1 and BRCA2 mutation carriers (United States). *Cancer Causes Control* 2003 Jun;14(5):439-46.
- Modugno F, Ness RB, Wheeler JE. Reproductive risk factors for epithelial ovarian cancer according to histologic type and invasiveness. *Ann Epidemiol*;2001 Nov; 11(8):568-74.
- Moher D, Cook DJ, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Quality of Reporting of Meta-analyses. *Lancet* 1999 Nov 27;354(9193):1896-900.
- Moher D, Schulz KF, Altman D. The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001 Apr 18;285(15):1987-91.
- Morrow AL, Ruiz-Palacios GM, Jiang X, et al. Human-milk glycans that inhibit pathogen binding protect breastfeeding infants against infectious diarrhea. *J Nutr* 2005 May;135(5):1304-7.
- Mortensen EL, Michaelsen KF, Sanders SA, et al. The association between duration of breastfeeding and adult intelligence.[see comment][erratum appears in JAMA 2002 Jun 12;287(22):2946]. *JAMA* 2002 May 8;287(18):2365-71.
- Ness RB, Grisso JA, Cottreau C, et al. Factors related to inflammation of the ovarian epithelium and risk of ovarian cancer.[see comment]. *Epidemiology* 2000 Mar;11(2):111-7.
- Neville MC, Keller RP, Seacat J, et al. Studies on human lactation. I. Within-feed and between-breast variation in selected components of human milk. *Am J Clin Nutr* 1984 Sep;40(3):635-46.
- Newburg DS, Peterson JA, Ruiz-Palacios GM, et al. Role of human-milk lactadherin in protection against symptomatic rotavirus infection. *Lancet* 1998 Apr 18;351(9110):1160-4.
- Newburg DS, Ruiz-Palacios GM, Morrow AL. Human milk glycans protect infants against enteric pathogens. *Annu Rev Nutr* 2005;25:37-58.
- Newcomb PA, Storer BE, Longnecker MP, et al. Lactation and a reduced risk of premenopausal breast cancer.[see comment]. *N Engl J Med* 1994 Jan 13;330(2):81-7.
- Nommsen LA, Lovelady CA, Heinig MJ, et al. Determinants of energy, protein, lipid, and lactose

concentrations in human milk during the first 12 mo of lactation: the DARLING Study. *Am J Clin Nutr* 1991 Feb;53(2):457-65.

Norris JM, Scott FW. A meta-analysis of infant diet and insulin-dependent diabetes mellitus: do biases play a role? *Epidemiology* 1996 Jan;7(1):87-92.

Nylen, O, Anderson, B, and Aniansson, G. Reduced frequency of acute otitis media in breast-fed infants. Abstract from the 5th International Congress of Pediatric Oto-Rhino-Laryngology Gent, Belgium, June 5-9. 1990. Ref Type: Abstract

O'Connor DL, Jacobs J, Hall R, et al. Growth and development of premature infants fed predominantly human milk, predominantly premature infant formula, or a combination of human milk and premature formula. *J Pediatr Gastroenterol Nutr* 2003 Oct;37(4):437-46.

O'Hara MW, Swain AM. Rates and risk of postpartum depression - a meta-analysis. *Int Rev Psychiatry* 1996;8(1):37-54.

Oddy WH, Holt PG, Sly PD, et al. Association between breast feeding and asthma in 6 year old children: findings of a prospective birth cohort study. *BMJ* 1999 Sep 25;319(7213):815-9.

Oddy WH, Kendall GE, Blair E, et al. Breastfeeding and cognitive development in children. *Adv Exp Med Biol* 2004;554:365-9.

Oddy WH, Kendall GE, Blair E, et al. Breast feeding and cognitive development in childhood: a prospective birth cohort study. *Paediatr Perinatal Epidemiol* 2003 Jan;17(1):81-90.

Ohlin A, Rossner S. Maternal body weight development after pregnancy. *Int J Obes* 1990 Feb;14(2):159-73.

Ohlin A, Rossner S. Factors related to body weight changes during and after pregnancy: the Stockholm Pregnancy and Weight Development Study. *Obes Res* 1996 May;4(3):271-6.

Olson CM, Strawderman MS, Hinton PS, et al. Gestational weight gain and postpartum behaviors associated with weight change from early pregnancy to 1 y postpartum. *Int J Obes Relat Metab Disord* 2003 Jan;27(1):117-27.

Owen CG, Martin RM, Whincup PH, et al. Effect of infant feeding on the risk of obesity across the life course: a quantitative review of published evidence. *Pediatrics* 2005 May;115(5):1367-77.

Owen CG, Martin RM, Whincup PH, et al. Does breastfeeding influence risk of type 2 diabetes in later life? A quantitative analysis of published evidence. *Am J Clin Nutr* 2006 Nov;84(5):1043-54.

Owen CG, Whincup PH, Gilg JA, et al. Effect of breast feeding in infancy on blood pressure in later life: systematic review and meta-analysis. [Review] [26 refs]. *BMJ* 2003 Nov 22;327(7425):1189-95.

Owen CG, Whincup PH, Odoki K, et al. Infant feeding and blood cholesterol: a study in adolescents and a systematic review.[see comment]. [Review] [88 refs]. *Pediatrics* 2002 Sep;110(3):597-608.

Pearson D, Kaur M, San P, et al. Recovery of pregnancy mediated bone loss during lactation. *Bone* 2004 Mar;34(3):570-8.

Pinelli J, Saigal S, Atkinson SA. Effect of breastmilk consumption on neurodevelopmental outcomes at 6 and 12 months of age in VLBW infants. *Adv Neonat Care* 2003 Apr;3(2):76-87.

Poiesz BJ, Ruscetti FW, Gazdar AF, et al. Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma. *Proc Natl Acad Sci U S A* 1980 Dec;77(12):7415-9.

Quigley MA, Cumberland P, Cowden JM, et al. How protective is breast feeding against diarrhoeal disease in infants in 1990s England? A case-control study. *Arch Dis Child* 2006 Mar;91(3):245-50.

Quinn PJ, O'Callaghan M, Williams GM, et al. The effect of breastfeeding on child development at 5 years: a cohort study. *J Paediatr Child Health* 2001 Oct;37(5):465-9.

Riman T, Dickman PW, Nilsson S, et al. Risk factors for invasive epithelial ovarian cancer: results from a Swedish case-control study. *Am J Epidemiol* 2002 Aug 15;156(4):363-73.

Risch HA, Marrett LD, Howe GR. Parity, contraception, infertility, and the risk of epithelial ovarian cancer. *Am J Epidemiol* 1994 Oct 1;140(7):585-97.

Risch HA, Weiss NS, Lyon JL, et al. Events of reproductive life and the incidence of epithelial ovarian cancer. *Am J Epidemiol* 1983 Feb;117(2):128-39.

Rønnestad A, Abrahamsen TG, Medbo S, et al. Late-onset septicemia in a Norwegian national cohort of extremely premature infants receiving very early full human milk feeding. *Pediatrics* 2005 Mar;115(3):e269-e276.

Rosenbaum PF, Buck GM, Brecher ML. Early child-care and preschool experiences and the risk of childhood acute lymphoblastic leukemia. *Am J Epidemiol* 2000 Dec 15;152(12):1136-44.

Rosenblatt KA, Thomas DB. Lactation and the risk of epithelial ovarian cancer. The WHO Collaborative Study of

- Neoplasia and Steroid Contraceptives. *Int J Epidemiol* 1993 Apr;22(2):192-7.
- Saarinén UM, Kajosaari M, Backman A, et al. Prolonged breast-feeding as prophylaxis for atopic disease. *Lancet* 1979 Jul 28;2(8135):163-6.
- Sassen ML, Brand R, Grote JJ. Breast-feeding and acute otitis media. *Am J Otolaryngol* 1994 Sep;15(5):351-7.
- Scariati PD, Grummer-Strawn LM, Fein SB. A longitudinal analysis of infant morbidity and the extent of breastfeeding in the United States. *Pediatrics* 1997 Jun;99(6):E5.
- Schanler RJ, Lau C, Hurst NM, et al. Randomized trial of donor human milk versus preterm formula as substitutes for mothers' own milk in the feeding of extremely premature infants. *Pediatrics* 2005 Aug;116(2):400-6.
- Schauberger CW, Rooney BL, Brimer LM. Factors that influence weight loss in the puerperium. *Obstet Gynecol* 1992 Mar;79(3):424-9.
- Schellscheidt J, Ott A, Jorch G. Epidemiological features of sudden infant death after a German intervention campaign in 1992. *Eur J Pediatr* 1997 Aug;156(8):655-60.
- Schultz Larsen F, Hanifin J. Epidemiology of Atopic Dermatitis. *Immunol Allergy Clin North Am* 2002;22(1):1-24.
- Sears MR, Greene JM, Willan AR, et al. Long-term relation between breastfeeding and development of atopy and asthma in children and young adults: a longitudinal study.[see comment]. *Lancet* 2002 Sep 21;360(9337):901-7.
- Seimyr L, Edhborg M, Lundh W, et al. In the shadow of maternal depressed mood: experiences of parenthood during the first year after childbirth. *J Psychosom Obstet Gynecol* 2004 Mar;25(1):23-34.
- Shay DK, Holman RC, Newman RD, et al. Bronchiolitis-associated hospitalizations among US children, 1980-1996. *JAMA* 1999;282(15):1440-6.
- Shu XO, Linet MS, Steinbuch M, et al. Breast-feeding and risk of childhood acute leukemia. *J Natl Cancer Inst* 1999 Oct 20;91(20):1765-72.
- Sichieri R, Field AE, Rich-Edwards J, et al. Prospective assessment of exclusive breastfeeding in relation to weight change in women. *Int J Obes Relat Metab Disord* 2003 Jul;27(7):815-20.
- Siskind V, Green A, Bain C, et al. Breastfeeding, menopause, and epithelial ovarian cancer. *Epidemiology* 1997 Mar;8(2):188-91.
- Smith DE, Lewis CE, Caveny JL, et al. Longitudinal changes in adiposity associated with pregnancy. The CARDIA Study. Coronary Artery Risk Development in Young Adults Study.[see comment]. *JAMA* 1994 Jun 8;271(22):1747-51.
- Smith MM, Durkin M, Hinton VJ, et al. Influence of breastfeeding on cognitive outcomes at age 6-8 years: follow-up of very low birth weight infants. *Am J Epidemiol* 2003 Dec 1;158(11):1075-82.
- Sowers M. Pregnancy and lactation as risk factors for subsequent bone loss and osteoporosis. *J Bone Miner Res* 1996 Aug;11(8):1052-60.
- Sowers M, Corton G, Shapiro B, et al. Changes in bone density with lactation.[see comment]. *JAMA* 1993 Jun 23;269(24):3130.
- Sowers MR, Clark MK, Hollis B, et al. Radial bone mineral density in pre- and perimenopausal women: a prospective study of rates and risk factors for loss. *J Bone Miner Res* 1992 Jun;7(6):647-57.
- Stamler J, Daviglus ML, Garside DB, et al. Relationship of baseline serum cholesterol levels in 3 large cohorts of younger men to long-term coronary, cardiovascular, and all-cause mortality and to longevity. *JAMA* 2000 Jul 19;284(3):311-8.
- Stenstrom C, Ingvarsson L. Otitis-prone children and controls: a study of possible predisposing factors. 1. Heredity, family background and perinatal period. *Acta Oto-Laryngologica* 1997 Jan;117(1):87-93.
- Strachan DP. Hay fever, hygiene, and household size. *BMJ* 1989 Nov 18;299(6710):1259-60.
- Stroup DF, Berlin JA, Morton SC, et al. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000 Apr 19;283(15):2008-12.
- Stuebe AM, Rich-Edwards JW, Willett WC, et al. Duration of lactation and incidence of type 2 diabetes. *JAMA* 2005 Nov 23;294(20):2601-10.
- Tai TY, Wang CY, Lin LL, et al. A case-control study on risk factors for Type 1 diabetes in Taipei City. *Diabetes Res Clin Pract* 1998 Dec;42(3):197-203.
- Taylor JS, Kacmar JE, Nothnagle M, et al. A systematic review of the literature associating breastfeeding with type 2 diabetes and gestational diabetes. *J Am Coll Nutr* 2005 Oct;24(5):320-6.
- Teale DW, Klein JO, Rosner B. Epidemiology of otitis media during the first seven years of life in children in

greater Boston: a prospective, cohort study. *J Infect Dis* 1989 Jul;160(1):83-94.

Titus-Ernstoff L, Perez K, Cramer DW, et al. Menstrual and reproductive factors in relation to ovarian cancer risk. *Br J Cancer* 2001 Mar 2;84(5):714-21.

Tung KH, Goodman MT, Wu AH, et al. Reproductive factors and epithelial ovarian cancer risk by histologic type: a multiethnic case-control study. *Am J Epidemiol* 2003 Oct 1;158(7):629-38.

Tung KH, Wilkens LR, Wu AH, et al. Effect of anovulation factors on pre- and postmenopausal ovarian cancer risk: revisiting the incessant ovulation hypothesis. *Am J Epidemiol* 2005 Feb 15;161(4):321-9.

Tyson JE, Lasky RE, Mize CE, et al. Growth, metabolic response, and development in very-low-birth-weight infants fed banked human milk or enriched formula. I. Neonatal findings. *J Pediatr* 1983 Jul;103(1):95-104.

U.S.Department of Health and Human Services. Healthy People 2010: Conference Edition. [II], 46-48. 2000. Washington, DC, U.S. Government Printing Office. Ref Type: Book, Whole

Uhari M, Mantysaari K, Niemela M. A meta-analytic review of the risk factors for acute otitis media.[see comment]. *Clin Infect Dis* 1996 Jun;22(6):1079-83.

UusiRasi K, Sievanen H, Pasanen M, et al. Association of physical activity and calcium intake with the maintenance of bone mass in premenopausal women. *Osteoporos Int* 2002 Mar;13(3):211-7.

Vennemann MM, Findeisen M, ButterfassBahloul T, et al. Modifiable risk factors for SIDS in Germany: results of GeSID. *Acta Paediatrica* 2005 Jun;94(6):655-60.

Vernacchio L, Lesko SM, Vezina RM, et al. Racial/ethnic disparities in the diagnosis of otitis media in infancy. *Int J Pediatr Otorhinolaryngol* 2004 Jun;68(6):795-804.

Virtanen SM, Knip M. Nutritional risk predictors of beta cell autoimmunity and type 1 diabetes at a young age 10046. *Am J Clin Nutr* 2003 Dec;78(6):1053-67.

Visalli N, Sebastiani L, Adorisio E, et al. Environmental risk factors for type 1 diabetes in Rome and province. *Arch Dis Child* 2003 Aug;88(8):695-8.

Vohr BR, Poindexter BB, Dusick AM, et al. Beneficial effects of breast milk in the neonatal intensive care unit on the developmental outcome of extremely low birth weight infants at 18 months of age. *Pediatrics* 2006 Jul;118(1):e115-e123.

Walker L, Freeland-Graves JH, Milani T, et al. Weight and behavioral and psychosocial factors among ethnically

diverse, low-income women after childbirth: II. Trends and correlates. *Women Health* 2004;40(2):19-34.

Warner R, Appleby L, Whitton A, et al. Demographic and obstetric risk factors for postnatal psychiatric morbidity. *Br J Psychiatry* 1996 May;168(5):607-11.

Wennergren G, Alm B, Oyen N, et al. The decline in the incidence of SIDS in Scandinavia and its relation to risk-intervention campaigns. Nordic Epidemiological SIDS Study. *Acta Paediatrica* 1997 Sep;86(9):963-8.

West RO. Epidemiologic study of malignancies of the ovaries. *Cancer* 1966 Jul;19(7):1001-7.

Whittemore AS, Harris R, Itnyre J. Characteristics relating to ovarian cancer risk: collaborative analysis of 12 US case-control studies. II. Invasive epithelial ovarian cancers in white women. Collaborative Ovarian Cancer Group.[see comment]. *Am J Epidemiol* 1992 Nov 15;136(10):1184-203.

Wigg NR, Tong S, McMichael AJ, et al. Does breastfeeding at six months predict cognitive development?[see comment]. *Aust N Z J Public Health* 1998 Apr;22(2):232-6.

Wright AL, Holberg CJ, Taussig LM, et al. Factors influencing the relation of infant feeding to asthma and recurrent wheeze in childhood.[see comment]. *Thorax* 2001 Mar;56(3):192-7.

Wynder EL, Dodo H, Barber HR. Epidemiology of cancer of the ovary. *Cancer* 1969 Feb;23(2):352-70.

Yen ML, Yen BL, Bai CH, et al. Risk factors for ovarian cancer in Taiwan: a case-control study in a low-incidence population. *Gynecol Oncol* 2003 May;89(2):318-24.

List of Acronyms and Abbreviations

<u>Abbreviation</u>	<u>Description</u>
Adj	Adjusted
AHRQ	Agency for Healthcare Research and Quality
ALL	Acute Lymphocytic Leukemia
AML	Acute myeloid leukemia
AOM	Acute Otitis Media
Bayley MDI	Bayley Mental Development Index
BF	Breastfeeding
BMC _{arm}	Forearm bone mineral content
BMD _{spine}	Lumbar spine bone mineral density
BMI	Body mass index
BW	Birth weight
CDC	Centers for Disease Control and Prevention
CES-D	Center for Epidemiological Studies Depression Scale
CI	Confidence interval
C-section	Delivery by cesarean section
CVD	Cardiovascular disease
DBP	Diastolic blood pressure
DIS	Diagnostic Interview Schedule
DM	Donor milk
DSM IV	Diagnostic Statistical Manual
Dx	Diagnosis
EPDS	Edinburgh Postnatal Depression Scale
Exclu	Exclusive
FH	Family history
GA	Gestational age
GDM	Gestational diabetes
H	Hospital or clinic controls
Hosp	Hospitalization
HRT	Hormone replacement therapy
Ht	Height
IHD	Ischemic heart disease
IOM	Institute of Medicine
IQ	Intelligence quotient
LDL	Low-density lipoprotein
LRTI	Lower respiratory track infection
LTC	Long-term recalls

<u>Abbreviation</u>	<u>Description</u>
MA	Meta-analysis
MANCOVA	Multivariate analysis of variance
MDI	Bayley Mental Development Index
MM	Exclusive mother's milk
Mo	Months
NA	Not applicable
ND	No data/not documented
NEC	Necrotizing Enterocolitis
NLSY79	National longitudinal survey of youth 1979
NS	Non-significant
OC	Ovarian cancer
OGTT	Oral glucose tolerance test
OR	Odds Ratio
P	Population controls
PA	Physical activity
PIAT	Peabody individual achievement test
PIAR	Poverty index ratio
PPVT-R	Peabody Picture Vocabulary Test
ROM	Recurrent Otitis Media
PF	Preterm formula
RDC	Research Diagnostic Criteria
RR	Relative risk
RTI	Respiratory infection
SBP	Systolic blood pressure
SDS	Standard deviation score
SE	Standard error
SES	Socioeconomic status
SR	Systematic review
SUDAAN	Software adjusted sample
TEP	Technical Expert Panel
Vit D suppl	Vitamin D supplementation
WISC-R	Weschler Intelligence Scale for Children
wk	Week
WPPSI-R	Wechsler Preschool and Primary Scales of Intelligence
Wt	Weight
Yr	Year

Appendix A. MEDLINE® Search Strategy

MEDLINE 1966-April 2006

#	Search History	Results
1	exp infant nutrition/	28620
2	exp Milk, Human/	10900
3	human milk.mp.	5231
4	(human adj2 milk).tw.	6040
5	breast milk.mp.	4994
6	breastmilk.mp.	315
7	breast feeding.mp.	18905
8	breastfeed\$.mp.	5099
9	breast fed.mp.	3434
10	breastfed.mp.	1339
11	(breast adj2 fed).tw.	3705
12	exp lactation/	24059
13	(lactating or lactation).mp.	33361
14	or/1-13	68554
15	exp HIV Infections/	150356
16	HIV.mp.	151368
17	*fatty acids/	21008
18	*amino acids/	32481
19	or/15-18	243664
20	14 not 19	66018
21	limit 20 to animals	32196
22	20 not 21	33822
23	limit 22 to english language	27054
24	follow-up studies/	308818
25	(follow-up or followup).tw.	340039
26	exp Case-Control Studies/	297283
27	(case adj20 control).tw.	44954
28	exp Longitudinal Studies/	508275
29	longitudinal.tw.	62652
30	exp Cohort Studies/	548011
31	cohort.tw.	73505
32	(random\$ or rct).tw.	324163
33	exp Randomized Controlled Trials/	39966
34	exp random allocation/	54141

35	exp Double-Blind Method/	84061
36	exp Single-Blind Method/	9446
37	randomized controlled trial.pt.	208988
38	clinical trial.pt.	420410
39	controlled clinical trials/	3005
40	(clin\$ adj trial\$.tw.	90860
41	((singl\$ or doubl\$ or trebl\$ or tripl\$) adj (blind\$ or mask\$)).tw.	80586
42	exp PLACEBOS/	24212
43	placebo\$.tw.	91878
44	exp Research Design/	197961
45	exp Evaluation Studies/	539139
46	exp Prospective Studies/	195111
47	exp Comparative Study/	1233790
48	or/24-47	2803651
49	23 and 48	8238
50	limit 49 to (addresses or bibliography or biography or case reports or congresses or consensus development conference or consensus development conference, nih or dictionary or directory or editorial or festschrift or government publications or interview or lectures or legal cases or legislation or letter or news or newspaper article or overall or patient education handout or periodical index)	360
51	49 not 50	7878

Appendix B. Sample Data Abstraction Forms

Evidence table for Systematic Reviews

Author, Year[UI]	Topic	
Systematic reviews/Meta-analyses/Both		
Databases searched (Dates of literature search)		
Countries where primary studies conducted (Developed countries only or mixed)		
Study design [No. of studies]		
No. of subjects		
Study population and sampling		
Intervention/Exposure		
Comparator		
Outcomes		
Methods used for meta-analyses		
Heterogeneity assessments		
Results		
Authors' conclusions		
Quality of the systematic review		
Comments		

Author, yr:

Topic of the systematic review*/MA:

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	
2	Types of exposure or intervention used	
3	Types of study designs used	
4	Study population	
	Reporting of Search strategy	
5	Search strategy, including time period and key words	
6	Effort to include all available studies (contact with authors)	
7	Databases and registries searched	
8	Method of addressing published articles other than English	
9	Method of handling abstracts and unpublished studies	
	Reporting of Methods	
10	Rationale for selection and/or coding of data	
11	Assessment of confounding	
12	Assessment of quality	
13	Assessment of heterogeneity	
14	Description of statistical methods sufficient to replicate	
15	Provision of appropriate tables and graphs	
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	
17	Tables giving descriptive information of each study included	
18	Results of sensitivity testing (subgroup analysis)	
19	Indication of statistical uncertainty of findings	
20	Quantitative assessment of bias (eg. publication bias)	
	Overall quality	

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	
8	Were results reported accurately (eg, no discrepancies between text and tables)?	
9	Were conclusions justified by the reported/collected data and analysis?	

Author, Year, UI #

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Pre-pregnancy BW (range): Pre-pregnancy BMI (range): Race: Enrolled/Evaluate: Location: Sites: Funding:					

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/confounders/limitations Comments			
			A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder		x	
			Blinding			
			Data collection		x	
			Withdraw and dropout			x
			Analyses		x	
			Intervention integrity			
			Overall:			

<i>Domain/question</i>	<i>Place an "X" in one</i>				
Selection Bias					
Are individuals selected to participate likely to be representative of target population?	Very likely	Somewhat likely	Not likely		
What % of selected individuals agreed to participate?	80-100	60-79	<60	ND	NA
Allocation Bias (RCTs only, for quasi-experimental, case-control/before/after, no control group or other skip to "Confounders")					
Is the method of random allocation stated	Yes	No			
If the method of random allocation is stated, is it appropriate	Yes	No			
Was the method of random allocation reported as concealed?	Yes	No			
Confounders					
Prior to the intervention, were there between group differences for important confounders reported in the paper?	Yes	No	Can't tell		
If there were differences between groups for important confounders, were they adequately managed in the analysis?	Yes	No	NA		
Were there important confounders NOT reported in the paper (describe above under quality score)?	Yes	No			
Blinding					
Was (were) the outcome assessor(s) blinded to the intervention or exposure status of the participants?	Yes	No	ND	NA	
Data Collection methods					
Were data collection tools shown or are they known to be valid?	Yes	No			
Were data collection tools shown or are they known to be reliable?	Yes	No			
Withdrawals and Dropouts					
Indicate the % of participants completing the study. (If the % differs by groups, record the lowest).	80-100	60-79	<60	ND	NA
Analysis					
Is there a sample size calculation or power calculation	Yes	Partially	No		
Is there a statistically significant difference between groups?	Yes	No	ND		
Are the statistical methods appropriate?	Yes	No	ND		
Indicate the unit of allocation	Community	Organization/group	Provider	Client	Institution
Indicate the unit of analysis	Community	Organization/group	Provider	Client	Institution
If the unit of allocation and analysis differed, was the cluster analysis done?	Yes	No	NA		
Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?	Yes	No	Can't tell		
Intervention Integrity					
What % of participants received the allocated intervention or exposure of interest?	80-100	60-79	<60	ND	NA
Was the consistency of the intervention measured (i.e. intervention was provided to all participants in the same way)?	Yes	No	ND	NA	
Is it likely that subjects received an unintended intervention (contamination or cointervention) that may influence the results?	Yes	No	Can't tell		

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Acute Otitis Media

Alho 1996

Duffy 1997

Sassen 1994

Stenstrom 1997

Uhari 1994 SRMA*

Vernacchio 2004

* Systematic Review/Meta-Analysis

Alho, 1996[8633605]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): ND Mean BW (range): ND % Male: ND Race: ND Enrolled/Evaluate: 2512/ 825 Location: Northern Finland Sites: Single Funding:	Retrospective Mean follow-up = 22 mo	All pregnant women in the two northernmost provinces of Finland with estimated dates of delivery between July1, 1985 and June30, 1986 were enrolled. The research program investigated the fetal period and the later development and illnesses of the children.	Breastfeeding among children older than 3 mo of age	Bottle feeding among children older than 3 mo of age

Alho, 1996[8633605]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Diagnostic criteria for AOM: At least one acute Symptom- Earache, fever, irritability, respiratory symptoms, Restless sleep, etc. and one pneumo-otoscopic finding- distinct redness and outward bulging or reduced mobility of the eardrum	Adjusted OR Confounders: Day care Parental smoking	OR=0.9 CI: 0.8 - 1.0	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder		x	
			Blinding			
			Data collection			x
			Withdraw and dropout			x
			Analyses		x	
			Intervention integrity		x	
			Overall: C High dropout rate			

Duffy, 1997 [9310540]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): Mean BW (range): % Male: 52 Race: 99% Caucasian Enrolled/Evaluate: 306/238 Location: USA Sites: Multi Funding: National Institute Of Child Health and Human Development Grant 19679	Prospective Follow-up duration: 24 mo	Consecutive infants at well baby visits shortly after birth. Infants with craniofacial abnormalities, genetic disorders, immune deficiencies were excluded	Changes in feeding modes were assessed at scheduled office visits - Exclusive milk breastfeeding - Combined breast- and formula- feeding Maternal guardians completed a questionnaire at 24months to determine the reliability of various postnatal parameters. K statistics: $r > 0.9$ for classification and duration of feeding modes (birth, 3, 6, and 12 mo of age)	Changes in feeding modes were assessed at scheduled office visits Formula feeding

Duffy, 1997 [9310540]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
AOM defined by presence of one or more of Fever, irritability, ear pain, pulling at the ears, And Tympanic membrane changes including Increased thickness, bulging, loss of landmarks, Decreased mobility	RR Confounder: age of colonization	Exclusive BF at 3 mo was associated with RR= 0.62; 95%CI: 0.43-.89; At 6 mo, RR=0.46; 95%CI: 0.29-0.74	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design	x		
			Confounder		x	
			Blinding			
			Data collection	x		
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity	x		
			Overall: B			

Sassen, 1994 [7978038]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): ND Mean BW (range): ND % Male: 53 Race: 96.6% Dutch Enrolled/Evaluate: 289/232 Location: Netherlands Sites: Multi; 1 urban and 2 rural Funding: ND	Prospective Mean follow-up = 23.6 mo	Children born between July 1987 and October 1988	Before breastfeeding was stopped (or per month)	Up to 4 mo after stopping

Sassen, 1994 [7978038]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
AOM diagnosis by physician; If purulent otorrhea; or if treatment for AOM given Criteria for AOM consisted of: Acute symptoms- ear ache, fever, irritability, restless sleep, etc. Otosopic signs- distinct redness and/or Outward bulging of the tympanic membrane Tympanometry results not used	Adjusted for 1. Number of siblings 2. Socio-economic status 3. Duration of breastfeeding	OR= 0.92 95% CI: 0.76-1.07	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design	x		
			Confounder		x	
			Blinding			
			Data collection		x	
			Withdraw and dropout			x
			Analyses	x		
			Intervention integrity	x		
			Overall: B Parents who are not Dutch could not Enroll their infants because the Questionnaire was not translated			

Stenstrom, 1997 [9039487]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): Mean BW (range): % Male: 61 Race: Enrolled/Evaluate: Location: Sweden Sites: Single/Multi Funding:	Case-control At examination children were between 3 and 7 years old Controls were between 4 and 6 years	All children born in the period 1978-81 registered to have had five or more episodes of AOM before age 30mo were selected and classified as cases (otitis-prone) From the official computerized population register in Malmo, 412 children with less than 5 episodes of AOM Were selected at random as a control group, and matched with otitis-prone children for age and sex so that each otitis-prone child had two matched controls	Ever breastfeeding	Never breastfeeding

Stenstrom, 1997 [9039487]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
AOM diagnosed otoscopically by an ENT physician, Pediatrician or general practitioner	Estimates not reported No adjustment for confounders	No differences between the groups found	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design		x	
			Confounder			x
			Blinding			
			Data collection		x	
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity		x	
			Overall: C No confounders adjusted			

Evidence table for Systematic Reviews of Breastfeeding and Acute otitis media

Author, Year[UI]	Topic	Uhari, 1996	Acute otitis media
Literature search (Dates)		Medline (1966-1994); Other databases searched? (yes); unpublished data used? (no)	
Countries where primary studies conducted		Finland, USA, Scotland	
Study eligibility / inclusion criteria		Search was limited to the English language and to human subjects. Only original studies of risk factors with adequate control groups and that reported actual numbers of patients were included	
Study design [No. Of studies]		2Case-control [12,17] 8Observational [[13,15,16,20,22,23,25,27]	
No. of subjects		Case control: 671 Observational: 4455 Breastfeeding > 3 months 2,548 Breastfeeding > 6 months 3,384 Breastfeeding (yes/no) 2, 193	
Study population (definition in included studies)		Studies of association between breastfeeding and acute otitis media and recurrence of AOM • Specific age groups of children in the cohorts are not specified	
Intervention/Exposure (definition in included studies)		The studies reviewed reported breastfeeding as yes/no or breastfeeding $\geq$ 3months vs < 3 months or breastfeeding $\geq$ 6 months vs < 6 months	
Comparator (definition in included studies)		Duration of breastfeeding	
Outcomes (definition in included studies)		10 studies evaluated the risk of Acute otitis media and recurrence of Acute otitis associated with breast milk feeding. The diagnosis of AOM varied, but pneumatic otoscopy was used in the diagnosis. There was no restriction on study inclusion according to the diagnostic criteria used. Twenty-two studies evaluating an array of risk factors for development and recurrence of AOM were included in the Meta-analysis. Seven of the twenty-two studies explicitly evaluated the association between breast milk feeding and AOM, while another 3/22 studies included breast milk feeding among other risk factors that were evaluated. All the studies agreed on the protective effect of breast milk feeding against AOM.	
Heterogeneity assessments		Pooled estimate of the relative risk and the 95% CI were calculated from data from studies that were sufficiently homogeneous. Pooled estimates of risks were derived with the random effects model. Details of clinical and statistical heterogeneity assessment is published elsewhere	
Quality assessments		Not done	
Publication bias assessments		Not done	
Statistical Analysis or meta-analytic methods		Random effects model used for the pooled estimates of risks	
Results		• breastfeeding for at least 3 months and AOM ;RR=0.87; 95% CI: 0.79 – 0.95; p = 0.003 • breastfeeding yes/no and recurrent AOM: RR=0.48; 95% CI: 0.32-0.72; p=0.0004 • breastfeeding $\geq$ 3 mo vs < 3 mo and recurrent AOM: RR= 0.69; 95% CI: 0.46-1.03; p=0.07 • breastfeeding $\geq$ 6 mo vs < 6 mo and recurrent AOM: RR= 0.69; 95% CI: 0.49-0.97; p=0.03	
Quality of the systematic review		Not done	
Author's interpretations of the results		• In the meta-analysis, breastfeeding was beneficial and breastfeeding even for only 3 months decreased the risk of AOM	

Author, Year[UI]	Topic	Uhari, 1996	Acute otitis media
		- Breastfeeding for 6 months or more decreased the recurrence of AOM - Breastfeeding should be promoted to protect against acute otitis media.	
Comments / Limitations		There was no restriction on inclusion according to the diagnostic criteria used for AOM. There are no details on how breastfeeding data was collected or duration of follow-up. Meta-analyses combined both case-control and cohort studies and did not address the potential biases in the individual studies.	

Author, yr: Uhari, 1994

Topic of the systematic review*/MA: Acute otitis media

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Yes
2	Types of exposure or intervention used	Yes
3	Types of study designs used	Yes
4	Study population	Yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Yes
6	Effort to include all available studies (contact with authors)	No (Medline only)
7	Databases and registries searched	Yes
8	Method of addressing published articles other than English	No
9	Method of handling abstracts and unpublished studies	No
	Reporting of Methods	
10	Rationale for selection and/or coding of data	No
11	Assessment of confounding	No
12	Assessment of quality	No
13	Assessment of heterogeneity	Yes
14	Description of statistical methods sufficient to replicate	Yes
15	Provision of appropriate tables and graphs	Yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Yes
17	Tables giving descriptive information of each study included	Partially
18	Results of sensitivity testing (subgroup analysis)	No
19	Indication of statistical uncertainty of findings	Yes
20	Quantitative assessment of bias (eg. publication bias)	No
	Overall quality	C

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Partially
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	Yes
5	Was individual study quality (such as sample size, study design, blinding, various	No

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
	biases and confounders, study subject attrition rate...etc.) assessed?	
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	No
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Yes
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Yes
8	Were results reported accurately (eg, no discrepancies between text and tables)?	N/A
9	Were conclusions justified by the reported/collected data and analysis?	Yes

Vernacchio, 2004 [15126021]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): Mean BW (range): 3,400,median % Male: 50.8 Race: 68.4% both parents white Enrolled/Evaluate: 15,113/11,349 Location: USA Sites: Multi Funding: Contract no. 1-HD-4-3221 National Institute of Child and Human Development and the National Institute on Deafness and other Communication Disorders	Prospective Followup- 6 mo	6 mo old infants who participated in the infant care practices study	Breastfeeding A child was considered to be breastfed at 6 mo if he or she was breastfeeding at the time of the 6 mo questionnaire, whether or not supplemental formula or solid foods were used	No breastfeeding A child was considered not to be breastfed at 6 mo if he or she was not breastfeeding at the time of the 6 mo questionnaire

Vernacchio,2004 [15126021]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
As part of the 6 mo questionnaire, each mother was asked health care provider in the month prior to date of completing the questionnaire. I so, they were asked to select a diagnosis from a menu of options that included "ear infection"	OR Confounders Gender Daycare attendance Mother's marital status Mother's age Mother's parity Number of children in the home	Unadjusted OR=0.66 95%CI: 0.59-0.74 Adjusted OR=0.69 95%CI: 0.61-0.78	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder		x	
			Blinding			
			Data collection	x		
			Withdraw and dropout	x		
			Analyses	x		
			Intervention integrity		x	
			Overall: B			

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Asthma

Burgess 2006

Gdalevich 2001 SRMA*

Kull 2004

Sears 2002

Wright 2001

* Systematic Review/Meta-Analysis

Burgess, 2006 [16585289]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 185 Mean BW (range): 3402 % Male: 52 Race: ND Enrolled/Evaluate: 7223/4964 Location: Australia Sites: Single Funding: ND	Prospective cohort with questionnaires at birth, 6 months, 5 and 14 years. Breastfeeding data collected at 6 months	Mothers and term singletons	Categories: Never breastfed < 3 weeks 3 – 6 weeks 7 weeks – 3 months ≥ 4 months	NA

Burgess, 2006 [16585289]

Outcome Definition	Statistical analyses and confounders adjusted	Results			Bias/limitations Comments				
At 14 years, mother reported asthma, qualitative answers to episodes of asthma past in 6 months. Supplemental questionnaire at 14 year follow-up requesting data on frequency of asthma meds, asthma-related sick days from school, asthma-related hospital admissions, parental history of asthma	Chi-square for categorical variables Logistic regression adjusting for breastfeeding, maternal asthma, paternal asthma, smoking early and late pregnancy, frequency of coughs and cold first 6 months, annual family income	Outcome	Breastfeeding duration			A: strong, B: moderate, C: weak	A	B	C
			No BF	3 wk-3 mo	≥4 mo				
		Asthma OR _{unadj} (95% CI)	1.0	1.03 (0.9-1.2)	1.03 (0.9-1.2)	Selection	x		
						Study design		x	
						Confounder		x	
						Blinding		x	
						Data collection		x	
						Withdraw and dropout			x
						Analyses			x
						Intervention integrity	NA		
		Of 4,964 subjects with breastfeeding and asthma data, 1408 (28%) mother-reported cases of adolescent asthma 3,720 returned 14 year follow-up questionnaire Nonsignificant relationship between duration of breastfeeding and asthma. Stratification for parental asthma or child's sex had no effect. No association for breastfeeding duration and asthma meds, asthma-related sick days from school or hospitalization			No reporting of adjusted OR, discrepancy for % males in table, dropout/withdraw > 30%, significant differences between completers and noncompleters, no definition or description of breastfeeding. Overall C				

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI] Topic	Gdalevich, 2001 Asthma
Literature search (Dates)	Medline (1966 to 1999); Other databases searched? (no); unpublished data used? (no)
Countries where primary studies conducted	Developed only: USA, Canada, UK, Italy, Australia and New Zealand
Study eligibility / inclusion criteria	Prospective studies; maternal recall of the child's feeding history of not more than 12 months; duration of BF for 3 months or more.
Study design [No. Of studies]	Prospective study design [12]
No. of subjects	8,183
Study population (definition in included studies)	Children (with and without family h/o atopy and asthma and combined population)
Intervention/Exposure (definition in included studies)	Exclusive breastfeeding (3 mo or more with no substitutes or supplements to mother's milk or solids)
Comparator (definition in included studies)	without ≥ 3 mo of exclusive breastfeeding
Outcomes (definition in included studies)	Asthma (Diagnosed by physician, consecutive reporting of wheezing beyond the age of 1 year, and treatment for lower respiratory tract allergy were defined as study outcomes)
Heterogeneity assessments	Yes by graphical presentation and application of heterogeneity test
Quality assessments	Unclear; studies critically appraised on 8 items (maternal recall; blinded to exposure; duration of bf; exclusivity of bf; diagnostic criteria; blinded ascertainment of outcome; age; control for confounding)
Publication bias assessments	Yes; calculation of the fail-safe N
Statistical Analysis or meta-analytic methods	Fixed effect model ; was also compared to random effects model Metric used odds ratio
Results	The summary OR was 0.47 (95% CI, 0.34-0.66) in the 5 studies (1788 subjects) with less than 2 years of follow-up, and 0.72 (95% CI, 0.62-0.84) in the 12 studies with 2 or more years of follow-up. The protective effect estimate of BF was more pronounced in the children with a family history of atopy in the subset of studies that assessed this population separately, OR= 0.52 (95% CI, 0.35-0.79), than in the studies in which both unstratified population data and children without atopic first-degree relatives were included, OR = 0.73 (95% CI, 0.62-0.86). When we further restricted the analysis to studies involving only children without a family history of atopy, the OR was 0.99 (95% CI, 0.48-2.03). Exclusion from the analysis of the studies with borderline blinding of diagnosing physicians ^{10,15,19} or those lacking a statement of BF exclusivity ²⁰ yielded respective ORs of 0.71 (95% CI, 0.61-0.84) and 0.72 (95% CI, 0.61-0.85).
Quality of the systematic review	A
Author's interpretations of the results	The summary analysis of the 12 prospective studies supports the role of exclusive BF in the prevention of childhood asthma. The protective effect was higher in the subgroup of children with a positive family history of asthma or atopy.
Comments / Limitations	Children age range not mentioned; included studies that did not adjust for confounding;

Author, yr: Gdalevich, 2001

Topic of the systematic review*/MA: Asthma

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Y
2	Types of exposure or intervention used	Y
3	Types of study designs used	Y
4	Study population	Y
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Y
6	Effort to include all available studies (contact with authors)	Y
7	Databases and registries searched	Y
8	Method of addressing published articles other than English	N
9	Method of handling abstracts and unpublished studies	Y
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Y
11	Assessment of confounding	Y
12	Assessment of quality	Y
13	Assessment of heterogeneity	Y
14	Description of statistical methods sufficient to replicate	Y
15	Provision of appropriate tables and graphs	Y
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Y
17	Tables giving descriptive information of each study included	Y
18	Results of sensitivity testing (subgroup analysis)	Y
19	Indication of statistical uncertainty of findings	Y
20	Quantitative assessment of bias (eg. publication bias)	N
	Overall quality	A

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Y
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Y
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Y
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	Y
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Y
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Y
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Y
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Y
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Y
9	Were conclusions justified by the reported/collected data and analysis?	Y

Kull, 2004; Kull, 2002
Wickman 2003

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): ND Mean BW (range): ND % Male: 50.5 Race: ND Enrolled/Evaluate: 4089/3601 Location: Sweden Sites: Multiple Funding: Nonprofit; government	Prospective, longitudinal cohort with final data collection at 4 year follow-up	All newborns from February 1994 through November 1996 in a predefined area of Stockholm, Sweden were invited to the study. 4089 (75%) of all infants were included. Data from all 4 questionnaires (at 2 months, 1, 2, and 4 years of age) were available for 3619 (88%) children. Complete answers on the subjects of breast-feeding, potential confounders, and outcome were required to be included in the analyses, leaving 3601 children (88% of the original study base) for analyses.	The independent variable of breast-feeding was categorized as follows: exclusive breast-feeding was dichotomized with the 25th percentile as the cutoff point (<4 months and ≥4 months) and as a 3-level categorical variable (0-2, 3-4, and ≥5 months). The variable for exclusive breast-feeding was used in combination with a dichotomized variable of partial breast-feeding (0-2 and ≥3 months) to disentangle the effects of exclusive breast-feeding and an additional period of partial breast-feeding. The period of additional partial breast-feeding was calculated from the point when exclusive breastfeeding was finished.	Exclusive breastfeeding 0-2 months Exclusive breastfeeding <4 months

Kull, 2004; Kull, 2002

Outcome Definition	Statistical analyses and confounders adjusted	Results			Bias/limitations Comments				
Asthma: At 4 years of age, asthma was defined as at least 4 episodes of wheezing during the last 12 months or at least 1 episode of wheezing during the same period if the child was receiving inhaled steroids. Early-onset persistent asthma implies that the child fulfilled the asthma criteria not only at 4 years but also during the first 2 years of life. Early-onset transient asthma denotes that the child was fulfilling asthma criteria during the first 2 years of life but not at 4 years of age. Late onset of asthma implies that the child was not classified as having asthma during the first 2 years of life but fulfilled the asthma criteria at 4 years.	The relationship between breast-feeding and health outcomes was analyzed with logistic regression, adjusted for maternal age, maternal smoking during pregnancy or at 2 months of age, and heredity. Heredity: Heredity for allergic diseases was defined as physician diagnosed asthma, hay fever, or both in combination with allergy to a furred pet, pollen, or both in 1 (single heredity) or 2 (double heredity) parents.	Outcome	Exclusive BF			A: strong, B: moderate, C: weak Selection Study design Confounder Blinding Data collection Withdraw and dropout Analyses Intervention integrity Overall: A	A	B	C
			0-2 mo	3-4 mo	≥5 mo				
		Asthma n/N OR _{adj} (95% CI)	48/541 1.0	41/701 0.67 (0.43-1.03)	111/2142 0.61 (0.42-0.86)				
		Asthma & No heredity n/N OR _{adj} (95% CI)	30/383 1.0	24/502 0.61 (0.36-1.06)	55/1500 0.48 (0.30-0.77)				
		Asthma & Heredity n/N OR _{adj} (95% CI)	18/158 1.0	17/199 0.76 (0.37-1.54)	56/642 0.81 (0.46-1.44)				
		Early-onset & persistent asthma n/N OR _{adj} (95% CI)	25/532 1.0	19/687 0.56 (0.30-1.04)	36/2118 0.35 (0.21-0.60)				
		Early-onset transient asthma n/N OR _{adj} (95% CI)	36/532 1.0	19/687 0.39 (0.22-0.69)	73/2118 0.51 (0.30-0.77)				
		Late onset of asthma n/N OR _{adj} (95% CI)	22/532 1.0	21/687 0.71 (0.38-1.31)	73/2118 0.82 (0.50-1.35)				
		Note: The interaction between asthma and heredity was not statistically significant.							
		The prevalence of asthma among children exclusively breastfed for less than 4 months was 9.1% compared with 6.4% among children breast-fed for 4 months or more (OR _{adj} , 0.72; 95% CI, 0.53-0.97). The OR _{adj} for asthma related to breast-feeding for 4 months or more was 0.58 (95% CI, 0.38-0.88) in children without heredity for allergic diseases and 0.73 (95% CI, 0.43-1.20) in children with heredity.							

Sears, 2002 [16585289]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): ND Mean BW (range):ND % Male: ND Race: ND Enrolled/Evaluate: 1037/1037 Location: New Zealand Sites: Single Funding: Government	Ambidirectional cohort study - enrolled at age 3 and followed prospectively up to 21 years. Accompanied and unaccompanied assessments at 3, 5, 7, 9, 11, 13, 15, 18, 21, and 26 years.	Population study of live-born children	Categories: Not breastfed Breastfed > 4 weeks Breastfed definition includes some formula feeding during birthing stay at hospital. Feeding history including duration of breastfeeding from regular, initially weekly, home & clinic visits during first 2-3 years	NA

Sears, 2002 [16585289]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
At age 9 years, ever having asthma from comprehensive questionnaires by interviewers with data on frequency of wheezing, asthma diagnosis, drugs, clinical characteristics, admissions; current asthma defined as positive response as reported by child or parent with accompanied symptoms of last 12 months; some verification by airway hyper-responsiveness	Chi-square for categorical variables Multivariate regression controlling for SES, birth order, sheepskin use in infancy, maternal smoking	Outcome	BF status		P value	A: strong, B: moderate, C: weak Selection x Study design x Confounder x Blinding x Data collection x Withdraw and dropout x Analyses x Intervention integrity N/A Overall: A	A	B	C
			No BF (n=504)	>4 weeks n=(533)					
		Asthma ever at 9 years OR _{adj} (95% CI)	n/N 27/417 (6%) 1.0	47/398 (12%) 1.93 (1.18-3.17)	0.0081				
		Current asthma at 26 years OR _{adj} (95% CI)	n/N 74/496 (15%) 1.0	113/484 (23%) 1.74 (1.26-2.40)	0.0008				
		Current asthma with AHR at 9 years OR _{adj} (95% CI)	n/N 11/409 (3%) 1.0	28/385 (7%) 2.83 (1.39-5.78)	0.0028				
		Asthma ever at 9 years Family history negative OR _{adj} (95% CI)	n/N 10/229 (4%)	23/216 (11%) 2.61 (1.21-5.62)					
		Asthma ever at 9 years Family history positive OR _{adj} (95% CI)	n/N 16/174 (9%)	23/174 (13%) 1.50 (0.77-2.96)	0.291				
		AHR, airway hyper-responsiveness to methacholine or salbutamol Multivariate analysis of current asthma at 9 years with breastfeeding > 4 weeks with adjusted OR 2.40 (1.36-4.26), p = 0.0027 Analysis for different duration of BF showed similar results of greater risk of asthma at 9 years for more							

Wright, 2000 [11065066]

Wright, 2001 [11182011]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): ND Mean BW (range): ND % Male: ND Race: ND Enrolled/Evaluate: 1,246/926 Location: USA Sites: Multiple Funding: Government	Prospective, longitudinal newborn cohort with questionnaires ascertaining respiratory health at 2, 3, 6, 9, 11, 13 years Additional data collection by health surveillance visit to MD	Healthy newborn infants	Children classified by duration of exclusive breast feeding: never breast fed, breast fed exclusively <4 months, breast fed exclusively >4 months	NA

Wright, 2000 [11065066]

Wright, 2001 [11182011]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Physician diagnosed asthma, wheezing or asthma symptoms reported ≥ 2 questionnaires from ages 6 to 13 years	Bivariate analyses between feeding history and asthma, and stratified by maternal asthma status Logistic regression to assess odds of asthma, recurrent wheeze were related to breast feeding and maternal asthma adjusting for confounders: maternal education, smoking status in 1 st year, sex, ethnicity, 2 or more siblings at home or day care use versus neither first 6 months, paternal asthma	Age 6-13: Nonsignificant for breastfeeding duration and asthma Nonsignificant for breastfeeding duration and asthma for children with non-asthmatic mothers OR _{adj} 2.1 (0.9-5.1) Significant for maternal asthmatics OR _{adj} 8.7 (3.4-22.2)	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design	x		
			Confounder	x		
			Blinding		x	
			Data collection		x	
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity	NA		
			Overall B No explanation for WD/dropouts, discrepancies in numbers for exclusive BF			

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Atopic Dermatitis

Gdalevich 2001 SRMA*

* Systematic Review/Meta-Analysis

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI]	Topic	Gdalevich 2001 (UI 11568741)	Atopic dermatitis
Literature search (Dates)		Medline (1966 to 5/2000); Other databases searched? No; unpublished data used? No	
Countries where primary studies conducted		Not reported; but all studies were performed in developed countries	
Study eligibility / inclusion criteria		Prospective studies of full-term infants in whom the ORs of atopic dermatitis associated with breastfeeding were reported or could be calculated. Exclusion: breastfeeding duration of <3 months; see paper for additional inclusion/exclusion criteria	
Study design [No. Of studies]		18 prospective studies were included	
No. of subjects		4158 subjects	
Study population (definition in included studies)		Full-term infants	
Intervention/Exposure (definition in included studies)		Exclusive breastfeeding for ≥ 3 months	
Comparator (definition in included studies)			
Outcomes (definition in included studies)		Strict diagnostic criteria for atopic dermatitis provided by the authors of the primary studies	
Heterogeneity assessments		Nonsignificant heterogeneity (P=0.27)	
Quality assessments			
Publication bias assessments			
Statistical Analysis or meta-analytic methods			
Results		Fixed effect, overall OR 0.68 (95% CI 0.52, 0.88); After exclusion of unblinded studies, OR 0.77 (95% CI 0.60, 0.98) Analyses restricted to those with positive family history, OR 0.58 (95%CI 0.41, 0.92) Analyses restricted to those without family history, OR 0.84 (95%CI 0.59, 1.19)	
Quality of the systematic review		A	
Author's interpretations of the results		There is a substantial protective effect of breastfeeding against atopic dermatitis in children with a family history of atopy.	
Comments / Limitations			

Author, yr: Gdalevich 2001

Topic of the systematic review*/MA: Atopic dermatitis

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Y
2	Types of exposure or intervention used	Y
3	Types of study designs used	Y
4	Study population	Y
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Y
6	Effort to include all available studies (contact with authors)	Y
7	Databases and registries searched	Y
8	Method of addressing published articles other than English	N
9	Method of handling abstracts and unpublished studies	N
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Y
11	Assessment of confounding	Y
12	Assessment of quality	Y
13	Assessment of heterogeneity	Y
14	Description of statistical methods sufficient to replicate	Y
15	Provision of appropriate tables and graphs	Y
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Y
17	Tables giving descriptive information of each study included	N
18	Results of sensitivity testing (subgroup analysis)	Y
19	Indication of statistical uncertainty of findings	Y
20	Quantitative assessment of bias (eg. publication bias)	N
	Overall quality	

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Y
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Y
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Y
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	Y
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Y
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Y
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Y
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Y/N
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Y
9	Were conclusions justified by the reported/collected data and analysis?	Y

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

CVD Mortality

Martin 2004 SRMA*

* Systematic Review/Meta-Analysis

Evidence table for Systematic Reviews of Breastfeeding

Author,Year[UI]	Topic	Martin, 2004	Cardiovascular Mortality
Literature search (Dates)		Medline (up to April 2003); Other databases searched? (yes); unpublished data used? (yes)	
Countries where primary studies conducted		Developed countries: US and UK	
Study eligibility / inclusion criteria		Articles were considered eligible for inclusion in the current review if infants who had been breastfed were compared with those who had been bottle (artificially) fed, if the outcome was cardiovascular disease or ischemic heart disease mortality, and if estimates of the association between having been breastfed in infancy and cardiovascular disease or ischemic heart disease mortality could be obtained from the paper or after correspondence with the authors.	
Study design [No. Of studies]		Historical cohort studies [4]	
No. of subjects		25,166 at baseline; 10,785 at follow-up	
Study population (definition in included studies)		Wingard, 1994: 1373 birth children in California Hertfordshire cohort: 5908 women and 10374 men born in 11 of 12 districts in Hertfordshire Boyd Orr: 4999 men and women from a survey of diet and health in pre-war Britain Caerphilly, 2003: 2512 meddle-aged men living in Caerphilly, South Wales	
Intervention/Exposure (definition in included studies)		- Any or exclusive breastfeeding. If results for both any or exclusive breastfeeding were presented in the paper, the exclusive breastfeeding association was used in the meta-analysis. - Prolonged breastfeeding: any or exclusive breastfeeding >1 year.	
Comparator (definition in included studies)		Exclusive bottle-feeding	
Outcomes (definition in included studies)		Cardiovascular disease or ischemic heart disease mortality	
Heterogeneity assessments		Cochran's Q statistic I ² test	
Quality assessments		No quality score was used, although specific aspects of the quality of each study, including control of confounding, loss to follow-up, recall bias, definition of breastfeeding, and sample size, were discussed	
Publication bias assessments		Not done, although there was a discussion on publication bias as one of the potential limitations of the meta-analyses.	
Statistical Analysis or meta-analytic methods		Random effect model	
Results		- All four studies were historical cohorts born between 1904 and 1939. - Random-effect models showed little difference in all cause mortality between breast and bottle-fed subjects (pooled rate ratio: 1.01; 95% CI: 0.91-1.13, p=0.8), and there was little evidence of heterogeneity. - Five observations from three studies suggested little or no association between breastfeeding and cardiovascular disease mortality in both males and females, and one suggested a possible adverse effect (Caerphilly cohort). In random effects meta-analysis, cardiovascular disease mortality was similar in breastfed versus bottlefed subjects (pooled rate ratio: 1.06; 95% CI: 0.94-1.20). There was no statistical evidence of between-study heterogeneity. - Ischemic heart disease mortality was 6% lower amongst males who had been breastfed in the Hertfordshire cohort, but 56% higher amongst breastfed females. This result is in line with point estimates from the Boyd Orr cohort suggesting that ischemic heart disease mortality was 10% lower amongst males who had been breastfed, but 40% higher amongst breastfed females (although there was little statistical evidence of interaction: p=0.2). In Caerphilly, however, ischemic heart disease mortality was 73% higher amongst breastfed males. In a random effects meta-analysis, there was evidence of heterogeneity.	

Author,Year[UI]	Topic	Martin, 2004	Cardiovascular Mortality
		• There was little evidence that prolonged breastfeeding was associated with all-cause mortality (pooled rate ratio: 0.94; 95% CI: 0.71–1.24), although there was moderate statistical evidence of heterogeneity. • There was weak evidence that prolonged breastfeeding was associated with a 16% increase (95% CI: 0.99–1.36; p=0.06) in cardiovascular disease mortality, and no evidence of inconsistency in estimates. • There was little evidence that prolonged breastfeeding was associated with ischemic heart disease mortality (rate ratio: 1.08; 95% CI: 0.88–1.31; p=0.5) and there was no heterogeneity.	
Quality of the systematic review		B	
Author's interpretations of the results		• There are at least three possible sources of bias in the studies reviewed: selection bias, information bias or recall bias (except for one study) and publication bias. Only 2 of the 4 studies had at least 70% of the target population in the follow-up. Recall bias is unlikely in only Hertfordshire cohort as breastfeeding was prospectively ascertained. Publication bias is possible but the studies were relatively large and unpublished estimates were obtained. • As confounding and bias may have distorted results from individual studies, the statistical combination of estimates into a combined rate ratio needs to be interpreted with caution. Our new analysis, together with evidence from published and unpublished literature, does not provide strong evidence that breastfeeding is related to all-cause or cardiovascular disease mortality. The confidence limits around our point estimates and the observed between-study heterogeneity for associations between breastfeeding and ischemic heart disease, however, do not rule out important beneficial or adverse cardiovascular effects of breastfeeding. • Although we found little evidence to support a cardioprotective effect of breastfeeding extending into old age, its beneficial influence on infant and child health and cognitive development supports the idea that it should be promoted as the infant feeding method of choice.	
Comments / Limitations		We agree with authors' conclusion and discussion on the limitations of the meta-analyses. However, given large heterogeneity across studies (especially for IHD mortality), meta-analyses might not be appropriate. At least for the outcome of IHD mortality, men and women should be combined separately because of apparent effect modification by gender. Results from Hertford cohort were only adjusted for age.	

Author, yr: Martin, 2004

Topic of the systematic review*/MA: CVD mortality

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Yes
2	Types of exposure or intervention used	Yes
3	Types of study designs used	Yes
4	Study population	Yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Yes
6	Effort to include all available studies (contact with authors)	Yes
7	Databases and registries searched	Yes
8	Method of addressing published articles other than English	No
9	Method of handling abstracts and unpublished studies	Yes
	Reporting of Methods	
10	Rationale for selection and/or coding of data	No
11	Assessment of confounding	Yes
12	Assessment of quality	Yes
13	Assessment of heterogeneity	Partially
14	Description of statistical methods sufficient to replicate	Yes
15	Provision of appropriate tables and graphs	Yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Yes
17	Tables giving descriptive information of each study included	Yes
18	Results of sensitivity testing (subgroup analysis)	No
19	Indication of statistical uncertainty of findings	Yes
20	Quantitative assessment of bias (eg. publication bias)	No
	Overall quality	B

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Yes
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	No
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Yes
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Partially
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	No
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Partially
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Yes
9	Were conclusions justified by the reported/collected data and analysis?	Yes

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Leukemia

Beral 2001 SRMA*

Davis 1998 SRMA*

Guise 2005 SRMA*

Kwan 2004 SRMA*

* Systematic Review/Meta-Analysis

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI] Topic	Beral (or USCCS Investigators), 2000; 2001	Leukemia
Literature search (Dates)	ND	
Countries where primary studies conducted	Mixed developed and developing countries	
Study eligibility / inclusion criteria	Not described. Meta-analyses were performed for the associations between breastfeeding and leukemia, Hodgkin's disease, non-Hodgkin's lymphoma, other childhood cancers and all childhood cancers combined. For the purpose of our review, we focus on childhood leukemia only.	
Study design [No. Of studies]	Case-control studies [15]	
No. of subjects	Cases: 7,401 Controls: 14,587	
Study population (definition in included studies)	Cases: Children with all leukemia, including ALL Controls: ND	
Intervention/Exposure (definition in included studies)	Ever breastfeeding Breastfeeding duration $\leq$ 6 months Breastfeeding duration $>$ 6 months	
Comparator (definition in included studies)	never breast fed	
Outcomes (definition in included studies)	All leukemia, including ALL	
Heterogeneity assessments	ND	
Quality assessments	ND	
Publication bias assessments	ND	
Statistical Analysis or meta-analytic methods	This is a primary study (UKCCS study) plus meta-analyses. No methods of meta-analyses were described.	
Results	The UKCCS results did not differ significantly from those of previous studies. Results from previous studies were combined with UKCCS results. For childhood leukemia, there is evidence of a statistically significant reduction in the OR associated with ever having been breastfed (OR=0.86, 95%CI 0.81-0.92) and having been breastfed for more than 6 months (OR=0.78, 95% CI 0.71-0.85).	
Quality of the systematic review	C	
Author's interpretations of the results	It is unclear whether the apparent small reduction in the risk of each type of childhood cancer reported here is a non-specific effect of breastfeeding on childhood cancer or whether it reflects some systematic bias shared by the majority of the case-control studies that have investigated etiological factors in childhood cancer. We are unable to decide which of these possibilities is the more likely.	
Comments / Limitations	This is a primary study (UKCCS study) plus meta-analyses. Authors focused on their results regarding the association between breastfeeding and childhood cancer in UKCCS study. No methods for the meta-analyses were described. They did not report the standard steps required for a systematic review, including explicit search criteria, inclusion and exclusion criteria, and quality assessment. However, we agree with the author's interpretations of the results.	

Author, yr: Beral 2000; 2001

Topic of the systematic review*/MA: Leukemia

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Yes
2	Types of exposure or intervention used	Yes
3	Types of study designs used	Yes
4	Study population	No
	Reporting of Search strategy	
5	Search strategy, including time period and key words	No
6	Effort to include all available studies (contact with authors)	No
7	Databases and registries searched	No
8	Method of addressing published articles other than English	No
9	Method of handling abstracts and unpublished studies	No
	Reporting of Methods	
10	Rationale for selection and/or coding of data	No
11	Assessment of confounding	No
12	Assessment of quality	No
13	Assessment of heterogeneity	Yes
14	Description of statistical methods sufficient to replicate	Yes
15	Provision of appropriate tables and graphs	Yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Yes
17	Tables giving descriptive information of each study included	No
18	Results of sensitivity testing (subgroup analysis)	No
19	Indication of statistical uncertainty of findings	Yes
20	Quantitative assessment of bias (eg. publication bias)	No
	Overall quality	C

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Unclear
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	No
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	No
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	No
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	No
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	No
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	No
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Yes
9	Were conclusions justified by the reported/collected data and analysis?	Yes

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI] Topic	Davis, 1998 Leukemia
Literature search (Dates)	Medline ("over the past 10 to 15 years"); Other databases searched? (no); unpublished data used? (no)
Countries where primary studies conducted	Primarily developed countries: UK, USA, Italy Included one developing country: China For the purpose of our review, we excluded the study from China
Study eligibility / inclusion criteria	Not described, except the author stated that hypothesis of the study was that artificial infant feeding (or never breast fed) increases the risk for cancer in childhood. For the purpose of our review, we focus on childhood leukemia only.
Study design [No. Of studies]	Case-control studies [5]
No. of subjects	All leukemia [reported in 1 study]- Cases: 171 Controls: 342 Acute lymphoblastic leukemia (ALL) [reported in all 5 studies]- Cases:1,356 Controls: 1,336 Acute non-lymphoblastic leukemia (ANLL) [reported in 2 studies]- Cases: 129 Controls: 379
Study population (definition in included studies)	Cases: Children with all leukemia, ALL or ANLL Controls: Children with no cancers, except for one study used children with rhabdomyosarcoma as controls
Intervention/Exposure (definition in included studies)	Any breastfeeding Long-term breastfeeding: more than 6 to 8 months
Comparator (definition in included studies)	Artificial infant feeding (or never breast fed)
Outcomes (definition in included studies)	"All leukemia, ALL, and ANLL"
Heterogeneity assessments	ND
Quality assessments	ND
Publication bias assessments	ND
Statistical Analysis or meta-analytic methods	No meta-analysis was performed. The odds ratio and 95% CI or p values for the relationship between infant feeding and cancer in childhood were provided in the papers; in other cases where the number of exposed and unexposed cases and controls were provided, additional effect estimates were calculated in order to compare studies.
Results	None of the included studies reported a significant association between all leukemia, ALL or ANLL and infant feeding.
Quality of the systematic review	C
Author's interpretations of the results	No evidence of an association between infant feeding and any childhood cancer except for Hodgkin's disease.
Comments / Limitations	The main purpose of this review was to examine the association between infant feeding and childhood cancer. Due to non-significant findings for leukemia, the author did not focus on the results on leukemia. Inclusion or exclusion criteria were not described, no assessments of study heterogeneity or methodology quality.

Author, yr: Davis, 1998

Topic of the systematic review*/MA: Leukemia

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Yes
2	Types of exposure or intervention used	Yes
3	Types of study designs used	Yes
4	Study population	Yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	No
6	Effort to include all available studies (contact with authors)	No
7	Databases and registries searched	Yes
8	Method of addressing published articles other than English	No
9	Method of handling abstracts and unpublished studies	No
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Yes
11	Assessment of confounding	No
12	Assessment of quality	No
13	Assessment of heterogeneity	No
14	Description of statistical methods sufficient to replicate	N/A
15	Provision of appropriate tables and graphs	Yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	No
17	Tables giving descriptive information of each study included	Yes
18	Results of sensitivity testing (subgroup analysis)	N/A
19	Indication of statistical uncertainty of findings	N/A
20	Quantitative assessment of bias (eg. publication bias)	N/A
	Overall quality	C

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	No
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	No
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	No
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	No
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	N/A
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	No
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Yes
9	Were conclusions justified by the reported/collected data and analysis?	Partially

Evidence table for Systematic Reviews of Breastfeeding

Author,Year[UI]	Topic	Guise, 2005	Leukemia
Literature search (Dates)		Medline (1990-March 2004); Other databases searched? (no); unpublished data used? (no)	
Countries where primary studies conducted		Developed countries only: Netherlands, Italy, US, Germany, Australia, New Zealand, Canada, England, Wales, Scotland, Sweden, France	
Study eligibility / inclusion criteria		Studies providing data regarding the association of breastfeeding and occurrence of childhood leukemia. A study must have been in full text, been published after 1990, contained more than 100 participants, had concurrent comparison groups, been conducted in a developed country, and been published in the English language.	
Study design [No. Of studies]		Case-control studies [10]	
No. of subjects		9,653 leukemia cases	
Study population (definition in included studies)		Cases: Childhood ALL or all childhood leukemias Controls: Matched or unmatched population (majority) and hospital controls	
Intervention/Exposure (definition in included studies)		Any measures of breastfeeding	
Comparator (definition in included studies)		Any	
Outcomes (definition in included studies)		Childhood ALL or all childhood leukemias defined in the original studies	
Heterogeneity assessments		ND	
Quality assessments		2 independent investigators rated study quality by using criteria from the US Preventive Services Task Force and the National Health Service Center for Reviews and Dissemination. Three levels of grade: good, fair and poor quality. Studies received a poor rating if the cases were not assessed reliably, if the groups assembled were not comparable, if there was considerable attrition or differences in nonrespondents between cases and controls, or if there was not adequate consideration given to confounding.	
Publication bias assessments		ND	
Statistical Analysis or meta-analytic methods		No meta-analysis or statistical analysis was performed.	
Results		Ten studies met all inclusion criteria and were reviewed for study quality to identify 2 good-quality case-control studies, 2 fair-quality studies, and 6 poor-quality studies. Of the 10 studies, 6 were conducted in European countries. All studies but 1 focused solely on childhood leukemia. The majority included <1000 cases. Notably, 6 of the studies explicitly sought to characterize the relationship between breastfeeding and leukemia as the primary objective, whereas the others included breastfeeding measures from the perspective measuring broader characteristics of the immune system and early infections in the etiology. The presented results focus on details and findings of the 2 good-quality studies, comparisons to the 2 fair quality studies, factors that distinguish them from studies rated as "poor" in quality, and implications for future research. The 2 good-quality studies (UKCCS and CCG studies) present conflicting results regarding the association between breastfeeding and leukemia. Similarly, the 2 fair-quality studies disagreed on the protective effect of breastfeeding.	
Quality of the systematic review		A	
Author's interpretations of the results		Of the 10 studies reviewed, only 4 were sufficient to provide at least fair-quality evidence regarding the association between maternal breastfeeding and childhood leukemia. Half of these 4 studies associated breastfeeding with a lower risk of ALL. Our review differs in methodology from previous meta-analyses performed by Beral et al and Kwan et al. Either of these groups examined studies for quality or used quality ratings as a determinate for	

Author,Year[UI]	Topic	Guise, 2005	Leukemia
		inclusion in analysis. The primary findings of our review indicate that there are few high-quality studies. The studies frequently failed to measure important factors such as breastfeeding exclusivity (reporting ever breastfed rather than quantifying breastfeeding by exclusivity combined with duration) and consideration of other important confounders such as SES and other infectious exposures (such as household or school contacts). None of the studies included in this systematic review were without flaw. An optimal study might be conducted within the framework of a large population-based registry or cohort with full access to medical records, pathology, and demographic data that would be able to accurately identify all cases of ALL diagnosed (with pathologic confirmation) at >1 year of age within a defined time period.	
Comments / Limitations		A meta-analysis by Beral, 2001 (or UKCCS investigators) seemed to have performed a comprehensive search, but they did not report the standard steps required for a systematic review, including explicit search criteria, inclusion and exclusion criteria, and quality assessment. This meta-analysis included 15 studies. Of these 15 studies, 8 were also included in Guise, 2005. Four were excluded because they were published before 1990, and 3 were conducted in developing countries. Guise, 2005 (10 studies included) and Kwan, 2004 (14 studies included) had same set of studies, except that Guise excluded studies published before 1990, and in developing countries. Guise, 2005 did not include non-peer-reviewed studies. Guise et al. discussed the differences between their review and previous meta-analyses by UKCCS investigators and Guise, 2005 in details. We thought their comments were fair. It would have been preferable if meta-analyses by taking into account study quality grading were also performed.	

Author, yr: Guise, 2005

Topic of the systematic review*/MA: Leukemia

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Yes
2	Types of exposure or intervention used	Yes
3	Types of study designs used	Yes
4	Study population	Yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Yes
6	Effort to include all available studies (contact with authors)	No
7	Databases and registries searched	Yes
8	Method of addressing published articles other than English	No
9	Method of handling abstracts and unpublished studies	No
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Yes
11	Assessment of confounding	Yes
12	Assessment of quality	Yes
13	Assessment of heterogeneity	N/A
14	Description of statistical methods sufficient to replicate	N/A
15	Provision of appropriate tables and graphs	Yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	No
17	Tables giving descriptive information of each study included	Yes
18	Results of sensitivity testing (subgroup analysis)	N/A
19	Indication of statistical uncertainty of findings	N/A
20	Quantitative assessment of bias (eg. publication bias)	N/A
	Overall quality	A

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Yes
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	Yes
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Yes
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Yes
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	No
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Yes
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Yes
9	Were conclusions justified by the reported/collected data and analysis?	Partially

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI]	Topic	Kwan, 2004	Leukemia
Literature search (Dates)		Medline (up to 2003); Other databases searched? (yes); unpublished data used? (yes)	
Countries where primary studies conducted		Mostly developed countries: Netherlands, Italy, US, Germany, Australia, New Zealand, Canada, England, Wales, Scotland, Sweden, France 2 developing countries: China, Russia,	
Study eligibility / inclusion criteria		Studies that presented data on any type of leukemia on children 15 years or younger in terms of an odds ratio and confidence interval and analyzed duration of breastfeeding in months were selected for the analysis.	
Study design [No. Of studies]		Case-control studies [n=14] *Note 12 studies in developed countries. Since this is a meta-analysis article, we could not separate these 12 studies. Three of the 14 studies were not peer-reviewed.	
No. of subjects		6,835 ALL cases and 1,216 acute myeloid leukemia (AML), which included ANLL	
Study population (definition in included studies)		Eight of the 14 studies excluded cases of leukemia in infants (usually children younger than 1 year of age) to avoid possible biases associated with premature cessation of breastfeeding in children with cancer and also because most leukemias occurring during infancy are known to have different etiologies from childhood leukemias.	
Intervention/Exposure (definition in included studies)		Short-term breastfeeding, defined as breastfeeding for six months or less Long-term breastfeeding, defined as breastfeeding for more than six months	
Comparator (definition in included studies)		No breastfeeding or never been breast-fed	
Outcomes (definition in included studies)		The ALL classification was generally straightforward, with only one study specifying "leukemia" instead of "ALL". Four studies used the classifications "other leukemias" or ANLL instead of AML. Nevertheless, they were included in the AML group since the majority of such cases are AML cases.	
Heterogeneity assessments		Chi-square statistic	
Quality assessments		ND, although there are comments on the potential biases for each included study in the summary tables	
Publication bias assessments		Funnel plot	
Statistical Analysis or meta-analytic methods		Fixed- and random-effect models. For each of the 14 articles reviewed, and R and its 95% CI were extracted. When available, ORs adjusted for SES were selected since SES was the most widely used potential confounder in these studies.	
Results		- No evidence of publication bias was apparent for studies reporting results on the association between long-term breastfeeding and risk of ALL ($p=0.58$) or AML ($p=0.46$). - A significant negative association was observed between short-term breastfeeding and ALL (OR=0.88, 95%CI 0.80-0.96), but the AML results (OR=0.90, 95%CI 0.80-1.02) were not significant. - A significant negative association was observed between long-term breastfeeding and ALL (OR=0.76, 95%CI 0.68-0.84), and AML (OR=0.85, 95%CI 0.73-0.98). - Heterogeneity statistics for leukemia types and duration of breastfeeding were not significant except for the ALL analysis addressing short-term breastfeeding ($p=0.03$). However, the results from fixed- and random-effect models were almost exactly the same. - The following table (modified from table 4 in the original table) shows the ORs for SES-adjusted data from the meta-analysis, as well as separately from the United Kingdom Childhood Cancer study (UKCCS) and the Children's Cancer Group (CCG) study. Overall, SES as a potential confounder appeared to play no substantial role in the findings of either the short-term or long-term breastfeeding studies.	

Author, Year[UI]	Topic	Kwan, 2004	Leukemia																																
		<table><tr><th rowspan="2">Duration of breastfeeding</th><th colspan="2">OR (95% CI)</th></tr><tr><th>ALL</th><th>AML</th></tr><tr><td>Meta-analysis</td><td>N=6470 (85%)</td><td>N=1179 (15%)</td></tr><tr><td>≤6 months</td><td>0.90 (0.82, 0.99)</td><td>0.91 (0.80, 1.04)</td></tr><tr><td>>6 months</td><td>0.75 (0.67, 0.85)</td><td>0.85 (0.73, 0.98)</td></tr><tr><td>UKCCS 2001</td><td>N=1401 (87%)</td><td>N=214 (13%)</td></tr><tr><td>≤6 months</td><td>0.90 (0.77, 1.04)</td><td>0.85 (0.60, 1.20)</td></tr><tr><td>>6 months</td><td>0.89 (0.75, 1.05)</td><td>0.65 (0.43, 1.00)</td></tr><tr><td>CCG study 1999</td><td>N=1744 (79%)</td><td>N=456 (21%)</td></tr><tr><td>≤6 months</td><td>0.86 (0.73, 1.01)</td><td>0.95 (0.68, 1.33)</td></tr><tr><td>>6 months</td><td>0.72 (0.60, 0.87)</td><td>0.57 (0.39, 0.84)</td></tr></table>		Duration of breastfeeding	OR (95% CI)		ALL	AML	Meta-analysis	N=6470 (85%)	N=1179 (15%)	≤6 months	0.90 (0.82, 0.99)	0.91 (0.80, 1.04)	>6 months	0.75 (0.67, 0.85)	0.85 (0.73, 0.98)	UKCCS 2001	N=1401 (87%)	N=214 (13%)	≤6 months	0.90 (0.77, 1.04)	0.85 (0.60, 1.20)	>6 months	0.89 (0.75, 1.05)	0.65 (0.43, 1.00)	CCG study 1999	N=1744 (79%)	N=456 (21%)	≤6 months	0.86 (0.73, 1.01)	0.95 (0.68, 1.33)	>6 months	0.72 (0.60, 0.87)	0.57 (0.39, 0.84)
Duration of breastfeeding	OR (95% CI)																																		
	ALL	AML																																	
Meta-analysis	N=6470 (85%)	N=1179 (15%)																																	
≤6 months	0.90 (0.82, 0.99)	0.91 (0.80, 1.04)																																	
>6 months	0.75 (0.67, 0.85)	0.85 (0.73, 0.98)																																	
UKCCS 2001	N=1401 (87%)	N=214 (13%)																																	
≤6 months	0.90 (0.77, 1.04)	0.85 (0.60, 1.20)																																	
>6 months	0.89 (0.75, 1.05)	0.65 (0.43, 1.00)																																	
CCG study 1999	N=1744 (79%)	N=456 (21%)																																	
≤6 months	0.86 (0.73, 1.01)	0.95 (0.68, 1.33)																																	
>6 months	0.72 (0.60, 0.87)	0.57 (0.39, 0.84)																																	
Quality of the systematic review		A																																	
Author's interpretations of the results		This meta-analysis demonstrated protective association between breastfeeding and risk of childhood ALL and possibly AML. However, three alternative explanations must be considered First, a systematic bias may be present in case-control studies of childhood leukemia arising from differential participation rates for case and control samples that differed in SES. In the current meta-analysis, participation rates could be calculated for only six of the 14 included studies. For five of these six studies, the participation rates for cases were higher than those for controls, and in most of these studies, the controls were higher SES than the cases. Therefore the association of breastfeeding with SES in combination with differential participation rates by SES could have biased the OR toward a protective effect. Second, the effect on ALL risk may be spurious, as suggested by the results of the two cohort studies (Ref#22 and 23), unfortunately, these studies provided no information regarding AML. Third, the protective effect of breastfeeding may not be limited to ALL. Further evaluation of the biological mechanisms while taking into consideration potential biases can be feasibly achieved with more large-scale case-control studies utilizing population-based, SES-matched controls. In conclusion, the potential protective effect of breastfeeding on risk of childhood ALL may be more complicated than the current literature suggests. Nevertheless, the available evidence suggests that such a protective effect exists for both ALL and AML, with the caveats noted.																																	
Comments / Limitations		We agree with authors' concerns regarding potential biases in the included studies. Their conclusion was appropriately cautious. However, the review still combined all studies regardless of these potential biases, and over-emphasized the importance of the results from the meta-analysis. We could not determine whether specific study should be included in the analyses. Perhaps combining results from only high-quality studies would be preferable, but unfortunately the review did not appraise the methodological quality of the primary studies.																																	

Author, yr: Kwan, 2004

Topic of the systematic review*/MA: Leukemia

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Yes
2	Types of exposure or intervention used	Yes
3	Types of study designs used	Yes
4	Study population	Yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Yes
6	Effort to include all available studies (contact with authors)	Yes
7	Databases and registries searched	Yes
8	Method of addressing published articles other than English	No
9	Method of handling abstracts and unpublished studies	Yes
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Yes
11	Assessment of confounding	Yes
12	Assessment of quality	Yes
13	Assessment of heterogeneity	Yes
14	Description of statistical methods sufficient to replicate	Yes
15	Provision of appropriate tables and graphs	Yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Yes
17	Tables giving descriptive information of each study included	Yes
18	Results of sensitivity testing (subgroup analysis)	Yes
19	Indication of statistical uncertainty of findings	Yes
20	Quantitative assessment of bias (eg. publication bias)	Yes
	Overall quality	A

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Yes
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	No
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Yes
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Yes
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Yes
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Yes
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Yes
9	Were conclusions justified by the reported/collected data and analysis?	Yes

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Cognitive

Agostoni 2001
Anderson 1999 SRMA*
Angelsen 2001
Der 2006
Drane 2000 SRMA*
GomezSanchiz 2003 2004
Jain 2002 SRMA*
Lawlor 2006
Mortensen 2002
Oddy 2003 2004
Quinn 2001

* Systematic Review/Meta-Analysis

Agostoni, 2001 [11787675]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): term Mean BW (range): % Male: 55% Race: Enrolled/Evaluate: 44/44 Location: Italy Sites: Single/Multi Funding:	Prospective cohort; Bayley results at 1 y/o were compared in breastfeeding ≥ 6 mo with 3-6 mo	Term infants; exclusively breastfed ≥ 3 mo		

Agostoni, 2001 [11787675]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Cognitive outcome of interest: Bayley Mental Development Index (MDI)	Adjustment for parity, maternal education, age, and smoking habits	After adjustment, Bayley MDI in 29 subjects breastfed >6 mo compared to 15 subjects breastfed 3-6 mo showed a 2.0 point advantage (95% CI -3.2, 7.3).	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design			x
			Confounder		x	
			Blinding		x	
			Data collection			
			Withdrawal and dropout			x
			Analyses			x
			Intervention integrity			
			Overall: C Small sample size, Bayley at young age			

Evidence table for Systematic Reviews

Author, Year[UI]	Topic	Anderson, 1999 [UI]	Cognitive Development
Systematic reviews/Meta-analyses/Both		Meta-analysis	
Databases searched (Dates of literature search)		Medline (1966-1996); references identified in bibliographies.	
Countries where primary studies conducted (Developed countries only or mixed)		Developed countries only: UK, US, Australia, Germany, New Zealand, and Spain	
Study design [No. of studies]		Observational (n=11)	
No. of subjects		7,081	
Study population and sampling		Studies of association between breastfeeding and cognitive development in full-term and low-birth weight infants (?premature)	
Intervention/Exposure		Many of the studies that were reviewed only reported breastfeeding versus bottle feeding, without more detailed information on exclusivity, frequency, or duration. Information on timing and duration of breastfeeding were included where available.	
Comparator			
Outcomes			
Methods used for meta-analyses		Both fixed effects and random effects were calculated	
Heterogeneity assessments		See below	
Results		- Unadjusted (fixed effects) pooled mean difference for the 11 observations was 5.32 (95% CI, 4.51, 6.14); heterogeneous across studies - Adjusted (fixed effects) pooled mean difference was 3.16 (95% CI, 2.35, 3.98); homogeneous across studies - An average adjusted benefit of 5.18 points was obtained for low-birth weight children across 6 available observations	
Authors' conclusions		After adjustment for key cofactors, breast-feeding was associated with significantly higher scores for cognitive development than was formula feeding	
Quality of the systematic review		B	
Comments		No appraisals of study quality (other than confounding); effect score is a standardized mean difference score of different cognitive measurements	

Author, yr: Anderson, 1999

Topic of the systematic review*/MA: Cognitive SR

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Y
2	Types of exposure or intervention used	Y
3	Types of study designs used	Y
4	Study population	Y
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Y/N
6	Effort to include all available studies (contact with authors)	N
7	Databases and registries searched	Y
8	Method of addressing published articles other than English	N
9	Method of handling abstracts and unpublished studies	N
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Y
11	Assessment of confounding	Y
12	Assessment of quality (in addition to confounding)	N
13	Assessment of heterogeneity	Y
14	Description of statistical methods sufficient to replicate	Y
15	Provision of appropriate tables and graphs	Y
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Y
17	Tables giving descriptive information of each study included	Y
18	Results of sensitivity testing (subgroup analysis)	Y
19	Indication of statistical uncertainty of findings	Y
20	Quantitative assessment of bias (eg. publication bias)	N
	Overall quality	

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Y
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Y
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Y/N
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	N
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	N
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Y/N
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Y
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Y
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Y
9	Were conclusions justified by the reported/collected data and analysis?	Y

Angelsen, 2001 [11517096]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 39.6 wk Mean BW (range): 3656 g % Male: Race: white Enrolled/Evaluate: 521/345/291 (5 yr) Location: Norway & Sweden Sites: Multi Funding:	Prospective cohort; 10% randomly selected from eligible population; Duration of breastfeeding: retrospectively recorded	White, parity 1 or 2, singleton, registered prior to 20 th wk gestation Excluded: <37 wk gestation, congenital malformation	<3 mo, 3-6 mo, ≥ 6 mo	

Angelsen, 2001 [11517096]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Cognitive outcomes of interest: Bayley mental scores at 13 mo Wechsler Preschool and Primary Scales of Intelligence (WPPSI-R) at 5 y/o	Maternal IQ, age, education, smoking	Bayley mental score at 13 mo: 117.7 (SD11.7) in ≥ 6 mo breastfeeding compared to 109.9 (SD13.1) in < 3 mo (P<0.001). There was a linear increase in mental development index plotted against breastfeeding duration (P<0.001). Maternal intelligence (Raven test score) was also related to duration of breastfeeding. Adjustment for differences in maternal intelligence reduced the OR of having a low mental developmental index among children who were breast fed for <3 mo to 1.6 (95%CI 1.1-2.3) from 3.2 (95%CI 1.7-5.9). Total IQ at 5 y/o: 111 (SD14.3) in ≥ 6 mo breastfeeding compared to 103.6 (SD14.6) in < 3 mo (P<0.001, Sheffe's test). Adjustment for differences in maternal intelligence reduced the OR of having a low IQ score among children who were breastfed for <3 mo to 1.5 (95%CI 1.0-2.1) from 2.8 (95%CI 1.4-5.3).	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design			x
			Confounder/bias		x	
			Blinding		x	
			Data collection		x	
			Withdrawal and dropout			x
			Analyses		x	
			Intervention integrity			
			Overall: B Results from multivariate analyses not reported.			

Der, 2006 UI 17020911

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): term Mean BW (range): % Male: Race: Enrolled/Evaluate: 3161 mothers and 5475 children Location: US Sites: Multi Funding:	Database analysis of a prospective study; sibling pairs analysis, and meta-analysis; Database from the US national longitudinal survey of youth 1979 (NLSY79) and children of the women in the survey, Peabody individual achievement test (PIAT) was administered to children between the ages of 5 and 14 biennially from 1986 to 2002, all outcomes standardized to a mean of 100 and SD of 15; maternal cognitive ability was measured with the Armed Forces Qualification Test (AFQT); For meta-analysis, only included studies that quantified the effect of breastfeeding status on cognitive ability after controlling for parental intelligence among full term infants (Medline 1966 to 1/2006 and other sources)	Excluded <35 wk gestation, <2500 g, born before 1979	Breastfeeding history obtained within a year of birth in most cases	

Der, 2006

Del, 2000

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Cognitive outcome of interest: Peabody individual achievement test	Adjustment for variables associated with breastfeeding in the survey, home environment (HOME-SF), child demographics, maternal characteristics	Unadjusted effect of breastfeeding +4.7 compared to non-breastfeeding (3161 mothers, 5475 children, 16,744 assessments); after adjustment for maternal AFQT score, education, age, family poverty, HOME stimulation score, and birth order, the difference became +0.52 (P=0.149) 332 pairs of sibling discordant for breastfeeding status and 545 discordant for duration of breastfeeding, difference between groups (status) = -0.63 (P=0.506); (duration) = -0.13 (P=0.866) Meta-regression of 9 unique studies (including the data from NLSY79): an advantage of breastfeeding of 0.16 after controlling for IQ and 8 additional confounders. Combined data from NLSY79 and sibling analysis study by Evenhouse (see separate extraction): estimate 0.025 (P=0.54) for breastfeeding status and 0.04 (P=0.271) for duration of breastfeeding	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design	x		
			Confounder	x		
			Blinding			
			Data collection			
			Withdraw and dropout	x		
			Analyses	x		
			Intervention integrity			
			Overall: A			
No details regarding breastfeeding history						

Domain/question	Place an "X" in one				
Selection Bias					
Are individuals selected to participate likely to be representative of target population?	Very likely	Somewhat likely	Not likely		
What % of selected individuals agreed to participate?	80-100	60-79	<60	ND	NA
Allocation Bias (RCTs only, for quasi-experimental, case-control/before/after, no control group or other skip to "Confounders")					
Is the method of random allocation stated	Yes	No			
If the method of random allocation is stated, is it appropriate	Yes	No			
Was the method of random allocation reported as concealed?	Yes	No			
Confounders					
Prior to the intervention, were there between group differences for important confounders reported in the paper?	Yes	No	Can't tell		
If there were differences between groups for important confounders, were they adequately managed in the analysis?	Yes	No	NA		
Were there important confounders NOT reported in the paper (describe above under quality score)?	Yes	No			
Blinding					
Was (were) the outcome assessor(s) blinded to the intervention or exposure status of the participants?	Yes	No	ND	NA	
Data Collection methods					
Were data collection tools shown or are they known to be valid?	Yes	No			
Were data collection tools shown or are they known to be reliable?	Yes	No			
Withdrawals and Dropouts					
Indicate the % of participants completing the study. (If the % differs by groups, record the lowest).	80-100	60-79	<60	ND	NA
Analysis					
Is there a sample size calculation or power calculation	Yes	Partially	No		
Is there a statistically significant difference between groups?	Yes	No	ND		
Are the statistical methods appropriate?	Yes	No	ND		
Indicate the unit of allocation	Community	Organization/group	Provider	Client	Institution
Indicate the unit of analysis	Community	Organization/group	Provider	Client	Institution
If the unit of allocation and analysis differed, was the cluster analysis done?	Yes	No	NA		
Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?	Yes	No	Can't tell		
Intervention Integrity					
What % of participants received the allocated intervention or exposure of interest?	80-100	60-79	<60	ND	NA
Was the consistency of the intervention measured (i.e. intervention was provided to all participants in the same way)?	Yes	No	ND	NA	
Is it likely that subjects received an unintended intervention (contamination or cointervention) that may influence the results?	Yes	No	Can't tell		

Evidence table for Systematic Reviews of Breastfeeding

Author,Year[UI]	Topic	Drane 2000	Cognitive Development
Literature search (Dates)		Database not specified; literature search from 1966-98	
Countries where primary studies conducted			
Study eligibility / inclusion criteria		Subjects born between 1960 and 1998, English language, studies examined some aspect of cognitive development	
Study design [No. Of studies]		Studies were assessed to determine if they met 3 methodological standards: definition of outcome: operational definition of cognitive outcome and outcome measure using standardized tests; correct classification of type of feeding: measure breast feeding as a continuous variable (duration of exclusive breast feeding or proportion of diet as breast milk) or at least as a 3-level categorical variable (exclusive breast or formula fed; partial breast fed); and control for potential confounding variables: socio-economic status, maternal education, birth weight, gestational age, birth order, gender in studies measuring verbal ability	
No. of subjects			
Study population (definition in included studies)		Term and preterm infants	
Intervention/Exposure (definition in included studies)			
Comparator (definition in included studies)			
Outcomes (definition in included studies)			
Heterogeneity assessments			
Quality assessments			
Publication bias assessments			
Statistical Analysis or meta-analytic methods			
Results		24/30 studies met entry criteria. 5 studies met the 3 methodological standards. Lucas 1992, Lucas 1994, Horwood 1998, Malloy 1998; Rogan 1993 (considered to have met 2.5 standards because the duration that a baby was mostly breast fed was measured) Lucas 1994 was a RCT, the remaining 5 studies were cohorts. Advantages in IQ as measured by the WISC-R, Bayley and McCarthy Scales were observed in Lucas 1992, Rogan 1993, Pollock 1994, and Horwood 1998. In term infants, effects on IQ in the range of 2-5 points (0.2-0.3 SD) were found. In low birth weight infants who received breast milk, Lucas 1992 reported an 8 IQ point advantage. Lucas 1994 did not find statistical significant difference in Bayley Mental Development Scale scores in infants fed solely on a diet of donor breast milk compared with infants fed solely on a diet of term formula. Malloy 1998 did not find an effect of infant feeding on WISC-R scores at 9-10 years of age in term infants.	
Quality of the systematic review		B	
Author's interpretations of the results		"The question of whether breast feeding and formula feeding have differential effects on cognitive development has not been comprehensively answered. A minority of the studies reviewed met acceptable standards for validity and, thus, are likely to provide inaccurate estimates of the association between infant feeding and cognitive development."	
Comments / Limitations		The methodological standards proposed may not be universally accepted. Study designs were not taken into consideration. Conclusions are also limited by the fact that almost all the studies were cohorts.	

Author, yr: Drane, 2000

Topic of the systematic review*/MA: cognitive

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Y
2	Types of exposure or intervention used	Y
3	Types of study designs used	Y
4	Study population	Y
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Y
6	Effort to include all available studies (contact with authors)	N
7	Databases and registries searched	N
8	Method of addressing published articles other than English	N
9	Method of handling abstracts and unpublished studies	N
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Y
11	Assessment of confounding	Y
12	Assessment of quality	Y
13	Assessment of heterogeneity	Y/N
14	Description of statistical methods sufficient to replicate	N/A
15	Provision of appropriate tables and graphs	Y
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Y
17	Tables giving descriptive information of each study included	Y/N
18	Results of sensitivity testing (subgroup analysis)	N/A
19	Indication of statistical uncertainty of findings	N/A
20	Quantitative assessment of bias (eg. publication bias)	N/A
	Overall quality	

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Y/N
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Y
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Y
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	N
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Y/N
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Y
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Y
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Y/N
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Y
9	Were conclusions justified by the reported/collected data and analysis?	Y

GomezSanchiz, 2003 [12635980]

GomezSanchiz, 2004 [15494884]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 39.6 wk Mean BW (range): 3344 g % Male: Race: Enrolled/Evaluate: 296/249 at 18 mo; 238 at 24 mo Location: Spain Sites: 1 rural, 1 urban Funding:	Prospective cohort from 1 rural (born 10/1995-9/1997) and 1 urban (born 3/1996-2/1998) sites, comparing breastfed > 4 mo, <4 mo, and formula fed; parents were informed of the project at 15 mo, Bayley was administered at 18 mo and 24 mo	37-42 wk gestation, Apgar at 5 minutes ≥ 7 , birthweight 2500-4500 g; Exclusion: multiple births, adopted babies, mechanical ventilation, hospitalization during the neonatal period, major physical malformation, chromosome disorder, psychosocial risks or any circumstances likely to result in "retarded development", height or weight <3%tile at 18 mo	Information about type and duration of feeding was taken from medical records	

GomezSanchiz, 2003 [12635980]

GomezSanchiz, 2004 [15494884]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Cognitive outcome of interest: Bayley Mental Development index	Adjusted for parental IQ	249/296 completed the study at 18 mo; 238 completed the study at 24 mo; the only difference between children taking part in the study and those not taking part was that the latter were less frequently first in birth order. Parental IQ was obtained only for 164 couples; their children had Mental Development Index 2.3 points higher than the children whose parents did not take part in IQ testing ($P < 0.05$). Infants were breastfed for a mean of 85.7 days $\pm$ SD 76.4 days. Duration of breastfeeding had a correlation with Mental Development Index at 18 mo ($r = 0.42$; $P < 0.001$) and at 24 mo ($r = 0.37$; $P < 0.001$). Mental Development Index at 18 mo Formula 106.0 $\pm$ 9.2 Breastfeeding ≤ 4 mo 111.5 $\pm$ 9.6 Breastfeeding > 4 mo 118.2 $\pm$ 8.9 $P < 0.05$ by Bonferroni in all post hoc pair comparisons At both 18 mo and 24 mo, after multiple linear regression adjusting for parental IQ, difference between formula and breastfeeding ≤ 4 mo no longer significant; difference of 4.3 points remained significant when comparing breastfeeding > 4 mo with ≤ 4 mo.	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design			x
			Confounder		x	
			Blinding		x	
			Data collection		x	
			Withdrawal and dropout		x	
			Analyses		x	
			Intervention integrity			
			Overall: B			

<i>Domain/question</i>	<i>Place an "X" in one</i>				
Selection Bias					
Are individuals selected to participate likely to be representative of target population?	Very likely	Somewhat likely	Not likely		
What % of selected individuals agreed to participate?	80-100	60-79	<60	ND	NA
Allocation Bias (RCTs only, for quasi-experimental, case-control/before/after, no control group or other skip to "Confounders")					
Is the method of random allocation stated	Yes	No			
If the method of random allocation is stated, is it appropriate	Yes	No			
Was the method of random allocation reported as concealed?	Yes	No			
Confounders					
Prior to the intervention, were there between group differences for important confounders reported in the paper?	Yes	No	Can't tell		
If there were differences between groups for important confounders, were they adequately managed in the analysis?	Yes	No	NA		
Were there important confounders NOT reported in the paper (describe above under quality score)?	Yes	No			
Blinding					
Was (were) the outcome assessor(s) blinded to the intervention or exposure status of the participants?	Yes	No	ND	NA	
Data Collection methods					
Were data collection tools shown or are they known to be valid?	Yes	No			
Were data collection tools shown or are they known to be reliable?	Yes	No			
Withdrawals and Dropouts					
Indicate the % of participants completing the study. (If the % differs by groups, record the lowest).	80-100	60-79	<60	ND	NA
Analysis					
Is there a sample size calculation or power calculation	Yes	Partially	No		
Is there a statistically significant difference between groups?	Yes	No	ND		
Are the statistical methods appropriate?	Yes	No	ND		
Indicate the unit of allocation	Community	Organization/group	Provider	Client	Institution
Indicate the unit of analysis	Community	Organization/group	Provider	Client	Institution
If the unit of allocation and analysis differed, was the cluster analysis done?	Yes	No	NA		
Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?	Yes	No	Can't tell		
Intervention Integrity					
What % of participants received the allocated intervention or exposure of interest?	80-100	60-79	<60	ND	NA
Was the consistency of the intervention measured (i.e. intervention was provided to all participants in the same way)?	Yes	No	ND	NA	
Is it likely that subjects received an unintended intervention (contamination or cointervention) that may influence the results?	Yes	No	Can't tell		

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI]	Topic	Jain, 2002	Cognitive Development
Literature search (Dates)		Medline (1996-2/2001); Other databases searched? (no); unpublished data used? (no)	
Countries where primary studies conducted			
Study eligibility / inclusion criteria		English language studies; independently assessed the relationship between breastfeeding and a cognitive outcome; tests of motor ability alone were not included	
Study design [No. Of studies]		30 birth cohorts; 2 RCTs; 5 school registry cohorts; 3 case-control studies	
No. of subjects		Ranged from 50 to >11,000	
Study population (definition in included studies)		Term and pre-term children	
Intervention/Exposure (definition in included studies)		Breastfeeding, breast milk, or choice to breast feed	
Comparator (definition in included studies)			
Outcomes (definition in included studies)		Appropriate outcome: standardized individual measure of general intelligence and that the assessment be done when the child was at least 2 years of age	
Heterogeneity assessments			
Quality assessments		Following were examined for each article: overall design of the study, sample size, target population, quality of feeding data, control of susceptibility bias, blinding, outcome measures, and format of results	
Publication bias assessments			
Statistical Analysis or meta-analytic methods			
Results		40 studies from 1929 to 2/2001 were included in this review 40 separate papers were published. Because some of these papers investigated the same sample, there were only studies of 33 different groups of children. 30 cohorts (27 with full-term children), 2 RCTs in preterm children, 5 school registry cohorts, and 3 case-control studies. 35 articles either studied mixed full-term and preterm infants, only full-term infants, or did not specify the gestational age or birth weight. 5 articles studied exclusively low birth weight infants. - Sample size ranged from 50 to >11,000 - Quality of feeding data (exclusive breastfeeding or not, timing of feeding data collection, source of feeding data, duration of breastfeeding): 9 articles met all 4 criteria (8 full-term cohorts); 15 articles did not adequately define breastfeeding by failing to report whether infants only received breast milk or were supplemented with formula or food. 2 preterm observational birth cohorts did not meet standard for appropriate definition of breastfeeding, it was defined as the "intent to breastfeed." 21 studies did not meet the standard for timing of data collection because the information was obtained either too late or too early. 27 included a feeding group for whom breastfeeding duration was at least 1 month. 9 studies controlled adequately for both socioeconomic status and level of stimulation of the child; 31 studies controlled for socioeconomic status. 15 studies stated that observers of the outcome were blind to feeding status. 22 studies used an appropriate measure of cognition. Of the remaining studies, 8 measured outcomes in children <2 years of age, 3 used screening measures to	

Author, Year[UI]	Topic	Jain, 2002	Cognitive Development
		assess cognition, 1 case control study used the presence of a learning disorder as its outcome measure, the other 2 case control studies used the diagnoses of pervasive developmental disorder and infantile autism, respectively, as outcome measures. 9. 33 studies reported some way to interpret the clinical significance of results, and 21 allowed calculation of an effect size. 10. 7 studies that found that the effects of breastfeeding statistically significant in the unadjusted analysis, became insignificant when controlling for socioeconomic status, stimulation, or other factors. 11. Only 2 studies met all the methodological standards. One study concluded that “any beneficial effect of breastfeeding on cognitive development is quite small in magnitude”, another study found that children who were breastfed had mean IQ scores 4.6 points higher than those never breastfed after controlling for socioeconomic status and other factors. Among the studies that controlled for socioeconomic status and stimulation/interaction of the child (not including the previous 2), 3 concluded that breastfeeding promotes cognitive development, and 4 did not.	
Quality of the systematic review		A	
Author's interpretations of the results		“No convincing evidence exists regarding the comparative effects of breastfeeding and artificial feeding on intelligence.”	
Comments / Limitations		Selection of methodological standards may not be acceptable to all investigators.	

Author, yr: Jain, 2002

Topic of the systematic review*/MA: cognitive

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Y
2	Types of exposure or intervention used	Y
3	Types of study designs used	Y
4	Study population	Y
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Y
6	Effort to include all available studies (contact with authors)	N
7	Databases and registries searched	Y
8	Method of addressing published articles other than English	Y
9	Method of handling abstracts and unpublished studies	N
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Y
11	Assessment of confounding	Y
12	Assessment of quality	Y
13	Assessment of heterogeneity	Y/N
14	Description of statistical methods sufficient to replicate	Y
15	Provision of appropriate tables and graphs	Y
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	N/A
17	Tables giving descriptive information of each study included	Y
18	Results of sensitivity testing (subgroup analysis)	N
19	Indication of statistical uncertainty of findings	Y/N

		Reporting yes/no
20	Quantitative assessment of bias (eg. publication bias)	N
	Overall quality	

*When grading a systematic review, please skip item #14, 18-20
Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Y
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Y
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Y
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	Y
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Y
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Y
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Y
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Y/N
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Y
9	Were conclusions justified by the reported/collected data and analysis?	Y

Lawlor 2006 [16466433]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): term and preterm (>98% term in followup N=3794) Mean BW (range): 3794 (% Male: 52% Race: Enrolled/Evaluate: 7223/3999 (5 yrs)/3794 (14 yrs) Location: Australia Sites: Single Funding:	Prospective cohort; Peabody Picture Vocabulary test at 5 years and Raven's standard progressive matrices at 14 years	Live birth who did not die before discharge from hospital	Duration of breastfeeding (never, <4 mos, ≥ 4 mos) was obtained from mothers at the 6-month followup assessment	

Lawlor 2006 [16466433]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Cognitive outcome of interest: Raven's standard progressive matrices at 14 years	Adjustment for sex, parental characteristics (maternal age, ethnicity, education, paternal education, family income, gravidity, maternal smoking), labor, Apgar scores, birthweight, height, BMI	At age 14 years, Never breastfed = 694, <4 months = 1372, ≥ 4 months = 1606 All parental characteristics were related to offspring IQ score. Unadjusted scores at age 14 showed a mean difference of 4.43 (95%CI 3.09 to 5.77) in <4 mos breastfeeding vs never breastfeeding; 8.20 (95%CI 6.89 to 9.49) in ≥ 4 mos breastfeeding vs never breastfeeding (P<0.001) Adjusted scores at age 14 (N=3099) for sex, parental characteristics (maternal age, ethnicity, education, paternal education, family income, gravidity, maternal smoking), labor, Apgar scores, birthweight, height, BMI showed a mean difference of 4.07 (95%CI 2.61 to 5.53) in <4 mos breastfeeding vs never breastfeeding; 6.79 (95%CI 5.33 to 8.26) in ≥ 4 mos breastfeeding vs never breastfeeding (P<0.001) Family income, parental education and breastfeeding explained 7.5% of the variation in intelligence at age 14. Loss to followup was selective, those subjects were more likely to have mothers who were from poorer social backgrounds, lower education, and younger; regression analysis repeated using Heckman's sample selection bias adjustment with maternal age, parental education, and family income as the selection variables; results of these regression models did not differ from those who had followup.	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design		x	
			Confounder	x		
			Blinding			x
			Data collection			
			Withdrawal and dropout			x
			Analyses	x		
			Intervention integrity			
			Overall: B No information on exclusivity of breastfeeding, unclear if cognitive evaluation is blinded, large drop out			

Domain/question	Place an "X" in one				
Selection Bias					
Are individuals selected to participate likely to be representative of target population?	Very likely	Somewhat likely	Not likely		
What % of selected individuals agreed to participate?	80-100	60-79	<60	ND	NA
Allocation Bias (RCTs only, for quasi-experimental, case-control/before/after, no control group or other skip to "Confounders")					
Is the method of random allocation stated	Yes	No			
If the method of random allocation is stated, is it appropriate	Yes	No			
Was the method of random allocation reported as concealed?	Yes	No			
Confounders					
Prior to the intervention, were there between group differences for important confounders reported in the paper?	Yes	No	Can't tell		
If there were differences between groups for important confounders, were they adequately managed in the analysis?	Yes	No	NA		
Were there important confounders NOT reported in the paper (describe above under quality score)?	Yes	No			
Blinding					
Was (were) the outcome assessor(s) blinded to the intervention or exposure status of the participants?	Yes	No	ND	NA	
Data Collection methods					
Were data collection tools shown or are they known to be valid?	Yes	No			
Were data collection tools shown or are they known to be reliable?	Yes	No			
Withdrawals and Dropouts					
Indicate the % of participants completing the study. (If the % differs by groups, record the lowest).	80-100	60-79	<60	ND	NA
Analysis					
Is there a sample size calculation or power calculation	Yes	Partially	No		
Is there a statistically significant difference between groups?	Yes	No	ND		
Are the statistical methods appropriate?	Yes	No	ND		
Indicate the unit of allocation	Community	Organization/group	Provider	Client	Institution
Indicate the unit of analysis	Community	Organization/group	Provider	Client	Institution
If the unit of allocation and analysis differed, was the cluster analysis done?	Yes	No	NA		
Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?	Yes	No	Can't tell		
Intervention Integrity					
What % of participants received the allocated intervention or exposure of interest?	80-100	60-79	<60	ND	NA
Was the consistency of the intervention measured (i.e. intervention was provided to all participants in the same way)?	Yes	No	ND	NA	
Is it likely that subjects received an unintended intervention (contamination or cointervention) that may influence the results?	Yes	No	Can't tell		

Mortensen 2002 [11988057]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): Mean BW (range): % Male: 490/973 (50% in cohort 1); 2280 (100% in cohort 2) Race: Enrolled/Evaluate: 973 had WAIS, 2280 had BPP test Location: Denmark Sites: Single/Multi Funding:	Prospective cohort but the breastfeeding information was collected retrospectively at the 1-year examination. There were 2 sub-cohorts; one participated in an ongoing developmental research program between 1982-1994 and took the WAIS at 27.2 years old; the other one took the draft board intelligence test: Borge Priens Prove (BPP) at 18.7 years; the BPP has a correlation of 0.82 with the full scale WAIS IQ.	Excluded twins		

Mortensen 2002 [11988057]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments							
Relationship of WAIS score and BPP score to duration of breastfeeding	Adjusted for parental social status and education, single mother status, mother's height, age, and weight gain during pregnancy, cigarette during 3 rd trimester, number of pregnancies, estimated gestational age, birthweight, birth length, and indices of pregnancy and delivery complications	Breastfeeding duration					A: strong, B: moderate, C: weak	A	B	C
							Selection		x	
							Study design		x	
							Confounder			x
							Blinding			x
							Data collection		x	
							Withdrawal and dropout		x	
							Analyses		x	
							Intervention integrity			
							Overall: B			

Domain/question	Place an "X" in one				
Selection Bias					
Are individuals selected to participate likely to be representative of target population?	Very likely	Somewhat likely	Not likely		
What % of selected individuals agreed to participate?	80-100	60-79	<60	ND	NA
Allocation Bias (RCTs only, for quasi-experimental, case-control/before/after, no control group or other skip to "Confounders")					
Is the method of random allocation stated	Yes	No			
If the method of random allocation is stated, is it appropriate	Yes	No			
Was the method of random allocation reported as concealed?	Yes	No			
Confounders					
Prior to the intervention, were there between group differences for important confounders reported in the paper?	Yes	No	Can't tell		
If there were differences between groups for important confounders, were they adequately managed in the analysis?	Yes	No	NA		
Were there important confounders NOT reported in the paper (describe above under quality score)?	Yes	No			
Blinding					
Was (were) the outcome assessor(s) blinded to the intervention or exposure status of the participants?	Yes	No	ND	NA	
Data Collection methods					
Were data collection tools shown or are they known to be valid?	Yes	No			
Were data collection tools shown or are they known to be reliable?	Yes	No			
Withdrawals and Dropouts					
Indicate the % of participants completing the study. (If the % differs by groups, record the lowest).	80-100	60-79	<60	ND	NA
Analysis					
Is there a sample size calculation or power calculation	Yes	Partially	No		
Is there a statistically significant difference between groups?	Yes	No	ND		
Are the statistical methods appropriate?	Yes	No	ND		
Indicate the unit of allocation	Community	Organization/group	Provider	Client	Institution
Indicate the unit of analysis	Community	Organization/group	Provider	Client	Institution
If the unit of allocation and analysis differed, was the cluster analysis done?	Yes	No	NA		
Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?	Yes	No	Can't tell		
Intervention Integrity					
What % of participants received the allocated intervention or exposure of interest?	80-100	60-79	<60	ND	NA
Was the consistency of the intervention measured (i.e. intervention was provided to all participants in the same way)?	Yes	No	ND	NA	
Is it likely that subjects received an unintended intervention (contamination or cointervention) that may influence the results?	Yes	No	Can't tell		

Oddy, 2003 [12562475] contains the same data as in Oddy, 2004 [15384602]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 37-39 wk Mean BW (range): % Male: Race: Enrolled/Evaluate: 2393 term at birth, 1444 at 6 years, 1371 at 8 years Location: Australia Sites: Single/Multi Funding:	Prospective cohort comparing no breastfeeding, <4 mo, 4-6 mo, and >4 mo of breastfeeding; Peabody Picture Vocabulary Test (PPVT-R) was administered at 6 y/o and a Performance subtest (Perceptual organization WISC- Block Design) at 8 y/o	Children with non-English speaking parents were excluded; preterm infants were also excluded		

Oddy, 2003 [12562475] contains the same data as in Oddy, 2004 [15384602]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Score of PPV-R at 6 y/o and Block Design score at 8 y/o	Association between breastfeeding duration and PPVT-R at 6 years and Block Design at 8 years were adjusted for gender, gestational age, maternal age and education, parental smoking, and the presence of older siblings.	90% of children were breastfed at some point, 28% were breastfed for >6 mo, solid foods were introduced at ≤6 mo in 88% of infants Both verbal IQ and performance scores increase with increasing maternal education combined with a longer duration of breastfeeding, with the most profound effect of breastfeeding occurring in the highest education groups (P<0.005). In the lower education groups, these trends were less consistent. After adjustment for covariates, there was an association between duration of breastfeeding and verbal IQ with a 3.56 point advantage for children breastfed >6 mo compared with children never breastfed (F=8.59, P=0.003). The adjusted association of full breastfeeding with the Performance subtest was not significant (F=1.49, P=0.223).	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder			x
			Blinding			x
			Data collection		x	
			Withdrawal and dropout			x
			Analyses		x	
			Intervention integrity			
			Overall: B			

Domain/question	Place an "X" in one				
Selection Bias					
Are individuals selected to participate likely to be representative of target population?	Very likely	Somewhat likely	Not likely		
What % of selected individuals agreed to participate?	80-100	60-79	<60	ND	NA
Allocation Bias (RCTs only, for quasi-experimental, case-control/before/after, no control group or other skip to "Confounders")					
Is the method of random allocation stated	Yes	No			
If the method of random allocation is stated, is it appropriate	Yes	No			
Was the method of random allocation reported as concealed?	Yes	No			
Confounders					
Prior to the intervention, were there between group differences for important confounders reported in the paper?	Yes	No	Can't tell		
If there were differences between groups for important confounders, were they adequately managed in the analysis?	Yes	No	NA		
Were there important confounders NOT reported in the paper (describe above under quality score)?	Yes	No			
Blinding					
Was (were) the outcome assessor(s) blinded to the intervention or exposure status of the participants?	Yes	No	ND	NA	
Data Collection methods					
Were data collection tools shown or are they known to be valid?	Yes	No			
Were data collection tools shown or are they known to be reliable?	Yes	No			
Withdrawals and Dropouts					
Indicate the % of participants completing the study. (If the % differs by groups, record the lowest).	80-100	60-79	<60	ND	NA
Analysis					
Is there a sample size calculation or power calculation	Yes	Partially	No		
Is there a statistically significant difference between groups?	Yes	No	ND		
Are the statistical methods appropriate?	Yes	No	ND		
Indicate the unit of allocation	Community	Organization/group	Provider	Client	Institution
Indicate the unit of analysis	Community	Organization/group	Provider	Client	Institution
If the unit of allocation and analysis differed, was the cluster analysis done?	Yes	No	NA		
Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?	Yes	No	Can't tell		
Intervention Integrity					
What % of participants received the allocated intervention or exposure of interest?	80-100	60-79	<60	ND	NA
Was the consistency of the intervention measured (i.e. intervention was provided to all participants in the same way)?	Yes	No	ND	NA	
Is it likely that subjects received an unintended intervention (contamination or cointervention) that may influence the results?	Yes	No	Can't tell		

Quinn 2001 [11885710]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): Mean BW (range): % Male: 53% Race: Enrolled/Evaluate: 4049/3880 Location: Australia Sites: Single Funding:	Prospective cohort; data were collected at enrollment, shortly after birth, at 6 months and at 5 years. Peabody Picture Vocabulary Test (PPVT-R) was administered at 5 years. Results were analyzed in relation to breastfeeding duration.	Singleton; At 5 years, children with major neurological abnormalities and those for whom the data were incomplete were excluded		

Quinn 2001 [11885710]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Score on PPVT-R	Adjusted for birthweight, poverty, maternal education, maternal age, time in daycare or preschool, number of children in the household at 5 years	Breastfeeding ≥ 6 mo 103.6 (SD 13.1) No breastfeeding 94.2 (SD 14.1) There was a significant trend towards increasing PPVT-R with increased duration of breastfeeding ($P=0.0000$) After adjustment, the mean for those breastfed ≥ 6 mo was 8.2 points (95%CI 6.5, 9.9) higher for females and 5.8 points (95%CI 4.1, 7.5) higher for males when compared to those never breastfed.	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design			x
			Confounder		x	
			Blinding			x
			Data collection		x	
			Withdrawal and dropout		x	
			Analyses		x	
			Intervention integrity			
			Overall: B			

Domain/question	Place an "X" in one				
Selection Bias					
Are individuals selected to participate likely to be representative of target population?	Very likely	Somewhat likely	Not likely		
What % of selected individuals agreed to participate?	80-100	60-79	<60	ND	NA
Allocation Bias (RCTs only, for quasi-experimental, case-control/before/after, no control group or other skip to "Confounders")					
Is the method of random allocation stated	Yes	No			
If the method of random allocation is stated, is it appropriate	Yes	No			
Was the method of random allocation reported as concealed?	Yes	No			
Confounders					
Prior to the intervention, were there between group differences for important confounders reported in the paper?	Yes	No	Can't tell		
If there were differences between groups for important confounders, were they adequately managed in the analysis?	Yes	No	NA		
Were there important confounders NOT reported in the paper (describe above under quality score)?	Yes	No			
Blinding					
Was (were) the outcome assessor(s) blinded to the intervention or exposure status of the participants?	Yes	No	ND	NA	
Data Collection methods					
Were data collection tools shown or are they known to be valid?	Yes	No			
Were data collection tools shown or are they known to be reliable?	Yes	No			
Withdrawals and Dropouts					
Indicate the % of participants completing the study. (If the % differs by groups, record the lowest).	80-100	60-79	<60	ND	NA
Analysis					
Is there a sample size calculation or power calculation	Yes	Partially	No		
Is there a statistically significant difference between groups?	Yes	No	ND		
Are the statistical methods appropriate?	Yes	No	ND		
Indicate the unit of allocation	Community	Organization/group	Provider	Client	Institution
Indicate the unit of analysis	Community	Organization/group	Provider	Client	Institution
If the unit of allocation and analysis differed, was the cluster analysis done?	Yes	No	NA		
Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?	Yes	No	Can't tell		
Intervention Integrity					
What % of participants received the allocated intervention or exposure of interest?	80-100	60-79	<60	ND	NA
Was the consistency of the intervention measured (i.e. intervention was provided to all participants in the same way)?	Yes	No	ND	NA	
Is it likely that subjects received an unintended intervention (contamination or cointervention) that may influence the results?	Yes	No	Can't tell		

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Gastrointestinal Infections

Chien 2001 SRMA*

Quigley 2006

* Systematic Review/Meta-Analysis

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI]	Chien, 2001 UI 11795054	Gastrointestinal (GI) Infection
Literature search (Dates)	Medline (1996 to 1/1998); Other databases searched? (Yes) Science Citation Index was used to identify frequently cited articles in the reference lists of studies from MEDLINE; unpublished data used? No	
Countries where primary studies conducted	Industrialized countries (UK, New Zealand, US, Australia, Scotland)	
Study eligibility / inclusion criteria	GI infection was defined to be any illness associated with vomiting, change in consistency or frequency of stools, or isolation of a known enteropathogenic bacterial or viral agent; GI outcome restricted to first year of life; excluded studies in which mortality was employed as a combined outcome event; excluded studies in which all the various types of infections were combined as an outcome measure	
Study design [No. Of studies]	Prospective (12 studies, 5473 subjects) and retrospective (2 studies, 504 subjects) cohort studies; case-control studies (2 studies, 331 pairs);	
No. of subjects	See above	
Study population (definition in included studies)		
Intervention/Exposure (definition in included studies)		
Comparator (definition in included studies)	Infant feeding practices grouped into either exclusive breastfeeding and partial/mixed feeding or exclusive artificial feeding	
Outcomes (definition in included studies)		
Heterogeneity assessments	Yes	
Quality assessments		
Publication bias assessments		
Statistical Analysis or meta-analytic methods	Fixed effect model; pooled results weighted according to sample size	
Results	Conflicting results on the effect of breastfeeding on GI infection: 9/16 studies (56%) yielded a statistically significant protective effect of breastfeeding on GI infections; Majority of studies suffered from methodological deficiencies; 4 studies fulfilled criteria of controlling for detection bias, analyses of confounders, having a clear definition of infant feeding practices and infectious outcomes; 3 of these studies reported breastfeeding was protective against GI infection; Pooled estimate of the cohort studies: OR 0.36 (95% CI 0.32, 0.41; heterogeneity $P < 0.01$) Pooled estimate of 2 case-control studies: OR 0.54 (95% CI 0.36, 0.80; heterogeneity $P = 0.35$)	
Quality of the systematic review	B	
Author's interpretations of the results	"No firm conclusion can be drawn about the magnitude of a protective effect of breastfeeding against infections in industrialized countries, there is no doubt that it is protective in the case of many infants. This trend is most obvious in the studies of the highest quality."	
Comments / Limitations	Random effect model would be more appropriate; weighted according to sample size; not analyzed for potential confounding	

Author, yr: Chien, 2001

Topic of the systematic review*/MA:

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Y
2	Types of exposure or intervention used	Y
3	Types of study designs used	Y
4	Study population	Y
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Y/N
6	Effort to include all available studies (contact with authors)	N
7	Databases and registries searched	Y
8	Method of addressing published articles other than English	N
9	Method of handling abstracts and unpublished studies	N
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Y
11	Assessment of confounding	Y/N
12	Assessment of quality	Y
13	Assessment of heterogeneity	Y
14	Description of statistical methods sufficient to replicate	Y
15	Provision of appropriate tables and graphs	Y
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Y
17	Tables giving descriptive information of each study included	Y
18	Results of sensitivity testing (subgroup analysis)	N
19	Indication of statistical uncertainty of findings	Y
20	Quantitative assessment of bias (eg. publication bias)	N
	Overall quality	

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Y/N
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Y
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Y
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	N
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Y/N
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Y
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Y/N
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Y/N
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Y
9	Were conclusions justified by the reported/collected data and analysis?	Y

Quigley, 2006 [16308409]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): Mean BW (range): % Male: 58 Race: Enrolled/Evaluate: 167 cases/ 137 controls Location: UK Sites: Multi Funding:	Case-control study; stratified by age group (0-3 mo, 3-5.9 mo, 6-8.9 mo, ≥ 9 mo), social deprivation score, location of practice (London, not London); questionnaire on current milk feeding	< 1 yr old and had data on infant feeding	Current exclusive breast milk (only milk received was breast milk, but many had been weaned on to solids), mixed feeding	Current formula

Quigley, 2006 [16308409]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
A diarrheal case was defined as “someone who presented to the General Practitioner with loose stools or significant vomiting < 2 weeks, in the absence of a known non-infectious cause and preceded by a symptom-free period of 3 weeks.”	age, sex, weaning, social class, travel, no access to food mixer, contact with person with diarrhea/vomiting in and outside household Multivariate analysis	Completed questionnaire: 190 cases (60%); 161 controls (90%); excluded 23 cases and 24 controls (no feeding data, no match, “other” feeding) Analyzed: 167 cases/ 137 controls Adjusted OR of diarrhea Current BF 0.36 (95% CI 0.18, 0.74) Current not BF 1 P=0.005 “Little evidence of the protection of breastfeeding persisting beyond 2 months following breastfeeding cessation.”	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder	x		
			Blinding			
			Data collection		x	
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity			
			Cannot rule out recall or response biases			

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Infant Mortality

Chen 2004

Chen 2004 [15121986]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): ND Mean BW (range): <2500g 26% live birth (controls) and 24% postneonatal death (cases) % Male: 50.3 live birth 59.8% postneonatal death Race: White 44.4 (controls) 50.0 (cases) Black 52.2 (controls) 46.6 (cases) Others 3.4 (controls) 3.4 (cases) Enrolled/Evaluate: NA Location: national level data Sites: Multi (population based) Funding: ND	Case control study design Follow-up NA	Subjects included in the 1988 US National Maternal and Infant Health Survey data to analyze the association between breastfeeding and postneonatal death.	Ever vs Never	Ever vs Never

Chen 2004 [15121986]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Post neonatal death: death after 28 days	Because of the oversampling of black and low birth weight infants SUDAAN adjusted analysis was performed to re-weight the sample for the overall estimates and to calculate the ORs and 95% confidence intervals in the final models. Logistic regression models Adjusted for mother's age, education, and smoking during pregnancy and infant's gender, race (except for race subgroup analyses), birth weight (except for birth weight subgroup analyses), congenital malformation reported at birth, live birth order, plurality, and WIC status.	Showed an OR of 0.79 (95% CI: 0.67–0.93) for ever breastfed. Race-specific analyses gave similar estimates for the OR, although the proportion ever breastfed was much lower in black infants. For duration of breastfeeding, comparing cases who survived 3 months ($n = 691$ in original sample and $n = 5363$ after adjustment with SUDAAN) and all controls, 3 months or more of breastfeeding showed an OR of 0.62 (95% CI: 0.46–0.82), less (ie, more protective) than the OR for ever/never breastfed (0.79; 95% CI: 0.67–0.93). The OR for overall postneonatal death was 0.74 (95% CI: 0.63–0.87). Among cases only, a proportional hazard model showed that the risk of death at any specific time was marginally lower in the ever breastfed infants (hazard ratio: 0.91; 95% CI: 0.79–1.06).	A: strong	A	B	C
			B: moderate			
			C: weak			
			Selection		x	
			Study design		x	
			Confounder		x	
			Blinding		x	
			Data collection		x	
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity		x	
			Overall B			

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Respiratory Tract Diseases

Bachrach 2003 SRMA*

* Systematic Review/Meta-Analysis

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI]	Topic	Bachrach, 2003 [12622672]	Respiratory Illness
Literature search (Dates)		Medline (1966-2002); Other databases searched? (yes); unpublished data used? (yes)	
Countries where primary studies conducted		Developed countries (US, Canada, New Zealand, Australia, Scotland, Norway)	
Study eligibility / inclusion criteria		Included studies from developed countries; healthy full-term infants; and provided data on exclusive breast feeding Excluded studies on cystic fibrosis and allergic conditions; sick, premature, and/or low birth weight infants	
Study design [No. Of studies]		Prospective cohort [5]; retrospective cohort [2]; ecological study [1]; cross-sectional [1]. Only cohort studies were combined.	
No. of subjects		3201 exposed / 1324 unexposed	
Study population (definition in included studies)		Healthy full term infants	
Intervention/Exposure (definition in included studies)		Exclusive breastfeeding at least of 2 months or 9 months of total (any) breastfeeding	
Comparator (definition in included studies)		None breastfeeding	
Outcomes (definition in included studies)		Hospitalization due to lower respiratory tract disease (LRTD) in infancy: bronchiolitis, asthma, pneumonia, empyema, and infections due to specific agents (eg, RSV)	
Heterogeneity assessments		Yes; No statistical heterogeneity. Used sensitivity analysis to test appropriateness of combining studies.	
Quality assessments		None	
Publication bias assessments		None (but the authors examined unpublished data)	
Statistical Analysis or meta-analytic methods		Random effects model; subgroup analyses for two covariates – smoking and SES	
Results		Summary relative risk from 7 cohort studies: 0.28 (0.14 - 0.54)	
Quality of the systematic review		A	
Author's interpretations of the results		Consistent effect was found among all primary studies that indicated a protective role of breastfeeding in LRTD	
Comments / Limitations		Included studies that had precise definition of breastfeeding	

Author, yr: Bachrach, 2003

Topic of the systematic review*/MA: Breastfeeding and LRTI

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Y
2	Types of exposure or intervention used	Y
3	Types of study designs used	Y
4	Study population	Y
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Y
6	Effort to include all available studies (contact with authors)	Y
7	Databases and registries searched	Y
8	Method of addressing published articles other than English	Y
9	Method of handling abstracts and unpublished studies	Y
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Y
11	Assessment of confounding	Y
12	Assessment of quality	N
13	Assessment of heterogeneity	Y
14	Description of statistical methods sufficient to replicate	Y
15	Provision of appropriate tables and graphs	Y
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Y
17	Tables giving descriptive information of each study included	Y
18	Results of sensitivity testing (subgroup analysis)	Y
19	Indication of statistical uncertainty of findings	N
20	Quantitative assessment of bias (eg. publication bias)	N
	Overall quality	A

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	partially
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	no
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	yes
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	yes
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	yes
8	Were results reported accurately (eg, no discrepancies between text and tables)?	yes
9	Were conclusions justified by the reported/collected data and analysis?	yes

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Obesity

Arenz 2004 SRMA*

Harder 2005 SRMA*

Owen 2005 SRMA*

* Systematic Review/Meta-Analysis

Evidence table for Systematic Reviews of Breastfeeding

Author,Year[UI]	Topic	Arenz, 2004 [15314625]	Obesity
Literature search (Dates)		Medline (1966-Dec 2003); Other databases searched? (yes); unpublished data used? (no)	
Countries where primary studies conducted		Developed countries: Germany, USA, Britain, Australia, New Zealand, Czech Republic	
Study eligibility / inclusion criteria		- Population-based cohort, cross-sectional or case-control studies with children older than 1 y at the last follow-up stage, published in English, French, Italian, Spanish or German. - Only studies with adjustment for at least three of the following relevant confounding or interacting factors birth weight, parental overweight, parental smoking, dietary factors, physical activity and socioeconomic status were included in this meta-analysis. - Odds ratio or relative risk had to be reported and age at the last follow-up had to be between 5 and 18 y; feeding-mode had to be assessed and reported and obesity as outcome had to be defined by body mass index (BMI) percentiles >90, 95 or 97 kg/m². If risk estimates were calculated for different percentile values in a particular study, the estimate for the highest percentile-value was included in the meta-analysis.	
Study design [No. Of studies]		Studies did not meet inclusion criteria for meta-analyses: prospective cohort study [11]; retrospective cohort [1]; cross-sectional study [4]; case-control study [2] Studies met inclusion criteria for meta-analyses: prospective cohort study [2]; cross-sectional study [7]	
No. of subjects		For studies met inclusion criteria for meta-analyses only: 69,000	
Study population (definition in included studies)		Children and adolescents. One study included some adult subjects but the various odds ratios were reported depending on age and not clear which odds ratio was used in the meta-analyses.	
Intervention/Exposure (definition in included studies)		Definition of feeding mode varied across studies: Never BF or partly BF < 3 months vs. BF ≥ 3 month; mostly or only BF vs. mostly or only formula feeding in the first 6 months; BF never vs. ever; BF never vs. > 6 months, BF groups: <1 week, 1 week-1 months, 2-3 months, 4-6 months, 7-9 month, > 9 months (exclusivity of BF not reported)	
Comparator (definition in included studies)			
Outcomes (definition in included studies)		Definition of childhood obesity varied across studies: BMI ≥97 th percentile, >95 th percentile, definition according to Cole et al.	
Heterogeneity assessments		Sensitivity analyses by testing the stability of the findings across different approaches in study design, exposure ascertainment and selection of study participants.	
Quality assessments		ND	
Publication bias assessments		Funnel plot	
Statistical Analysis or meta-analytic methods		Fixed-effect and random-effect model: both crude and adjusted odds ratios (AOR) of the studies were used.	
Results		- In total, 28 studies met the inclusion criteria for the systematic review, 19 of them were not eligible for the meta-analysis. - The pooled crude OR for breast-feeding and obesity defined as BMI >90th, 95th or 97th percentile could be calculated for six studies. In the fixed-effects model, the OR was 0.67, 95% CI (0.62, 0.73). In the random effects model an almost identical OR was found (data not shown). - The AOR calculated for nine studies was 0.78, 95% CI (0.71, 0.85) for both fixed and random-effects model. - Sensitivity analyses were performed, comparing the studies according to the following criteria: cohort study or cross-sectional study, different definitions of breast-feeding, different definitions of obesity, different age-groups and number of potential confounders considered for adjustment. The protective effect of breast-feeding was more pronounced in studies with adjustment for less than seven potential confounding factors compared to adjustment for seven or more potential confounding factors. - The funnel plot showed an asymmetric pattern, which was due to a particular study. The funnel plot regression analysis did not reject the null hypothesis of symmetry (df=8,	

Author,Year[UI]	Topic	Arenz, 2004 [15314625]	Obesity
		P=0.71), suggesting that there was no publication bias.	
Quality of the systematic review		A	
Author's interpretations of the results		"The results from meta-analysis indicate that breast-feeding is associated with a small but consistent protective effect against obesity risk in later childhood."	
Comments / Limitations		Highly heterogeneous definitions of breastfeeding across studies overall meta-analysis was inappropriate, but sensitivity analyses were performed to reduce heterogeneity. Sensitivity analyses are very important but very few studies in each stratum due to small number of studies met the criteria for meta-analyses.	

Author, yr: Arenz, 2004

Topic of the systematic review*/MA:

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Yes
2	Types of exposure or intervention used	Yes
3	Types of study designs used	Yes
4	Study population	Yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Yes
6	Effort to include all available studies (contact with authors)	Yes
7	Databases and registries searched	Yes
8	Method of addressing published articles other than English	No
9	Method of handling abstracts and unpublished studies	No
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Yes
11	Assessment of confounding	Yes
12	Assessment of quality	Yes
13	Assessment of heterogeneity	Yes
14	Description of statistical methods sufficient to replicate	Yes
15	Provision of appropriate tables and graphs	Yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Yes
17	Tables giving descriptive information of each study included	Yes
18	Results of sensitivity testing (subgroup analysis)	Yes
19	Indication of statistical uncertainty of findings	Yes
20	Quantitative assessment of bias (eg. publication bias)	Yes
	Overall quality	A

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Yes
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	ND
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Yes
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Yes
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Yes
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Yes
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Partially
9	Were conclusions justified by the reported/collected data and analysis?	Yes

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI]	Topic	Harder, 2005 [16076830]	Obesity
Literature search (Dates)		Medline (1966-December 2003); Other databases searched? (yes); unpublished data used? (no)	
Countries where primary studies conducted		Developed countries: UK, US, Canada, Germany, Australia, New Zealand, Czechoslovakia	
Study eligibility / inclusion criteria		1) an original report comparing breastfed subjects with exclusively formula-fed subjects (referent group) of any given age, 2) report the odds ratio and 95% CI (or data to calculate them) of overweight or obesity associated with breastfeeding, and 3) report the duration of breastfeeding for at least one exposure group.	
Study design [No. Of studies]		Cohort (including cross-sectional design) [16]; case-control study [1]	
No. of subjects		Ranged 66 to 32,200 total 120,831	
Study population (definition in included studies)		Primarily children. Two studies included adults.	
Intervention/Exposure (definition in included studies)		Median duration of breastfeeding categories: < 1 month (reference), 1-3 month, 4-6 months, 7-9 months, > 9month Each month increase in the duration of breastfeeding: We calculated a study specific regression coefficient and corresponding 95% CI for each study by use of a log-linear model.	
Comparator (definition in included studies)		The resulting odds ratio and 95% CI for change in risk for each month of breastfeeding were pooled with a random-effect model	
Outcomes (definition in included studies)		Any definition of overweight or obesity, varied across studies	
Heterogeneity assessments		ND	
Quality assessments		ND	
Publication bias assessments		Funnel plot with Begg test and Egger test	
Statistical Analysis or meta-analytic methods		Unadjusted odds ratios and 95% CI were calculated directly from the data given in the articles, where possible. Otherwise, the published odds ratio and 95% CI were used. Three approaches were taken: First, duration of breastfeeding as the independent variable and the weighted odds ratio for overweight in breastfed probands, compared with formula-fed subjects, as the dependent variable (for meta-regression). Second, the random-effect model pooled odds ratio for overweight in breastfed subjects was calculated separately for five predefined categories of duration of breastfeeding. Third, the pool-first method was used to combine the regression coefficients obtained from the studies (for trend estimation). For meta-regression analysis, all duration-specific odds ratios had to be related to the respective duration of breastfeeding. The median of the upper and lower limits of each category of duration of breastfeeding was assigned to the particular estimate in each study. Estimates were plotted against the respective duration of breastfeeding as independent variable. A weighted meta-regression with duration of breastfeeding as the covariate was perform (random-effects model). For studies that provided data for more than two categories of duration of breastfeeding, we applied the "pool-first method" to quantify the dose-response relation. This was possible for 11 studies. We calculated a study specific regression coefficient and corresponding 95% CI for each study by use of a log-linear model. The resulting odds ratio and 95% CI for change in risk for each month of breastfeeding were pooled with a random-effect model.	
Results		From the 17 studies that reported duration of breastfeeding, 14 gave data for more than one category of duration of breastfeeding, leading to 52 estimates included in the meta-regression analysis. In the weighted meta-regression, duration of breastfeeding was significantly negatively related to risk of overweight (regression coefficient: 0.94, 95%CI 0.89-0.98).	

Author, Year[UI]	Topic	Harder, 2005 [16076830]	Obesity																														
		- From 1 month of breastfeeding onward, the risk of subsequent overweight continuously decreased up to a reduction of more than 30%, reaching a plateau at 9 months of breastfeeding. <table><tr><td></td><th colspan="5">Duration of breastfeeding</th></tr><tr><td></td><th><1 mo</th><th>1-3 mo</th><th>4-6 mo</th><th>7-9 mo</th><th>>9 mo</th></tr><tr><td>No. of duration-specific study estimates</td><td>5</td><td>14</td><td>15</td><td>11</td><td>7</td></tr><tr><td>OR for overweight</td><td>1.0</td><td>0.81</td><td>0.76</td><td>0.67</td><td>0.68</td></tr><tr><td>95% CI</td><td>0.65, 1.55</td><td>0.74, 0.88</td><td>0.67, 0.86</td><td>0.55, 0.82</td><td>0.50, 0.91</td></tr></table> - Each month of breastfeeding was found to be associated with a 4% decrease in risk (OR: 0.96/month of breastfeeding, 95%CI 0.94-0.98). A fixed-effects model revealed a similar pooled OR and a nearly identical 95% CI (OR: 0.96/month of breastfeeding, 95%CI 0.95-0.98) - In only two of these studies was the influence of the duration of exclusive breastfeeding analyzed. The pooled OR for risk of overweight per month of exclusive breastfeeding was 0.94 (95%CI 0.89-0.99, random-effect model) - Subgroup analyses revealed that the definition of overweight influenced the estimate only slightly. In studies that used BMI to define overweight, the pooled OR was 0.96 (95%CI 0.94-0.98) for 8 studies, while the OR was 0.93 (95%CI 0.87-0.99) for the 3 studies that used another measure to define overweight or obesity. - The age at examination had only a marginal influence on the magnitude of the effect of duration of breastfeeding on risk of overweight. The pooled OR from all 5 studies investigating probands up to or including 5 years of age was 0.97 (95%CI 0.94-0.99), while in older subjects aged 6 or more years, it was 0.96 (95%CI 0.93-0.99) for 6 studies. - No evidence of publication bias was observed, as indicated by a symmetric funnel plot and a nonsignificant Begg test and Egger test.		Duration of breastfeeding						<1 mo	1-3 mo	4-6 mo	7-9 mo	>9 mo	No. of duration-specific study estimates	5	14	15	11	7	OR for overweight	1.0	0.81	0.76	0.67	0.68	95% CI	0.65, 1.55	0.74, 0.88	0.67, 0.86	0.55, 0.82	0.50, 0.91	
	Duration of breastfeeding																																
	<1 mo	1-3 mo	4-6 mo	7-9 mo	>9 mo																												
No. of duration-specific study estimates	5	14	15	11	7																												
OR for overweight	1.0	0.81	0.76	0.67	0.68																												
95% CI	0.65, 1.55	0.74, 0.88	0.67, 0.86	0.55, 0.82	0.50, 0.91																												
Quality of the systematic review		B																															
Author's interpretations of the results		"Using three different techniques, we show that a longer duration of breastfeeding is associated with a larger decrease in risk of overweight. Each of the three methods used in our study has its own advantages and limitations." "In summary, we found that the duration of breastfeeding is inversely and linearly associated with the risk of overweight. The risk of overweight was reduced by 4% for each month of breastfeeding. This effect lasted up to a duration of breastfeeding of 9 months and was independent of the definition of overweight and age at follow-up. Even if interpreted as being of relatively small size, this association, if causal, might be of importance for the general population. Since the majority of studies analyzed here used partially breastfed subjects, it might be concluded that, beyond exclusive breastfeeding, also longer partial breastfeeding up to 9 months leads to a greater decrease in risk of overweight in later life, which might be considered in future clinical recommendation."																															
Comments / Limitations		The meta-analyses addressing the relationship between duration of breastfeeding and the risk of overweight was well-performed. However, the authors' conclusion implied a causal relationship, which is not appropriate because most of the "cohort" studies included in the analyses were in fact cross-sectional studies because the breastfeeding exposure was ascertained retrospectively at the time of the assessment of obesity. Causality cannot be ascertained in cross-sectional design. Most importantly, crude OR was used in the meta-analyses, which is the major limitation.																															

Author, yr: Harder, 2005

Topic of the systematic review*/MA: Obesity

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Yes
2	Types of exposure or intervention used	Yes
3	Types of study designs used	Yes
4	Study population	Yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Yes
6	Effort to include all available studies (contact with authors)	Yes
7	Databases and registries searched	Yes
8	Method of addressing published articles other than English	No
9	Method of handling abstracts and unpublished studies	No
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Yes
11	Assessment of confounding	Partially
12	Assessment of quality	Partially
13	Assessment of heterogeneity	No
14	Description of statistical methods sufficient to replicate	Yes
15	Provision of appropriate tables and graphs	Yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Yes
17	Tables giving descriptive information of each study included	Yes
18	Results of sensitivity testing (subgroup analysis)	Yes
19	Indication of statistical uncertainty of findings	Yes
20	Quantitative assessment of bias (eg. publication bias)	Yes
	Overall quality	B

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Yes
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	ND
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Partially
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Yes
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Yes
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Partially
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Yes
9	Were conclusions justified by the reported/collected data and analysis?	Yes

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI] Topic	Owen, 2005 Obesity
Literature search (Dates)	Medline (completed in September 2003); Other databases searched? (yes); unpublished data used? (no)
Countries where primary studies conducted	Developed and developing countries: UK, Germany, USA, Canada, Italy, Turkey, Australia, New Zealand, Slovak Republic, China, Czech Republic, Sweden
Study eligibility / inclusion criteria	61 studies that compared a measure of obesity (quantitatively or narratively) among breastfed and formula-fed subjects were considered. Studies that defined the odds of obesity or being overweight for breastfed and formula-fed subjects were reported more often and were included in a meta-analysis; 28 studies with 29 estimates (1 study reported results for 2 populations) met this inclusion criterion.
Study design [No. Of studies]	Historical cohort, prospective cohort, cross-sectional studies: total 61 studies, and 28 studies included in the meta-analyses
No. of subjects	N=298,900 in the overall meta-analysis
Study population (definition in included studies)	28 studies (298,900 subjects) provided 29 unadjusted odds ratios relating the initial infant feeding method and obesity. Four observations were for infants, 23 for children, and 2 for adults.
Intervention/Exposure (definition in included studies)	Breastfed vs. formula-fed. The exclusiveness of infant feeding was based on the classification given in each article. The feeding groups were defined as being mutually exclusive in 4 studies, the breastfed group included mixed feeders in 7 studies and the formula-fed group included mixed feeders in 7 studies. In 2 studies in which infants were breastfed exclusively, the exclusivity of formula feeding could not be determined.
Comparator (definition in included studies)	
Outcomes (definition in included studies)	Any measures of obesity or adiposity. Most studies used a percentile cutoff based on BMI, describing subjects at the tail of the distribution. The 95th or 97th percentile was used most often, although some studies used cutoff values as low as the 85th percentile. A smaller number of studies used absolute BMI values or cutoff values based on standardized weight or weight for height. Studies involving infants used definitions based on percentiles of weight for length or percentiles of weight only.
Heterogeneity assessments	Q test with chi-square statistics
Quality assessments	ND
Publication bias assessments	Begg funnel plot. Reporting bias was also assessed.
Statistical Analysis or meta-analytic methods	Results from fixed-effects models are reported throughout, because these reflect only the random error within each study, are more conservative because they are less affected by results of smaller studies that show stronger associations, and make no assumptions about the representativeness of the available studies. Meta-regression was used to examine the influence of the following factors (defined a priori) with a test for trend: study size (<500, 500–2500, or >2500 subjects), age group at outcome measurement (infants <=1 year of age, young children >1 to 9 years of age, older children 10 to <16 years of age, or adults >=16 years of age), year of birth, and response rates (analyzed as a continuous variable). The effects of adjustment for factors such as parental body size (mostly BMI), socioeconomic status, and maternal smoking were examined in 6 studies that provided data before and after adjustment for all 3 of these factors. The effects of study methods, particularly the method of ascertainment of infant feeding status (whether contemporary or recalled over a period of >=3 years), study response rate, and definition of obesity (equivalent to <95th percentile, 95th to <97th percentile, or >=97th percentile of BMI), were examined with meta-regression and sensitivity analyses.
Results	Among 61 observational studies that reported on the effects of infant feeding on a measure of adiposity in later life, 28 studies (298 900 subjects) provided 29 unadjusted odds ratios relating the initial infant feeding method and obesity. Four observations were for infants, 23 for children, and 2 for adults.

Author, Year[UI]	Topic	Owen, 2005	Obesity
		• In a fixed-effects model including all studies, breastfed subjects were less likely to be defined as obese than were formula-fed subjects (OR: 0.87; 95% CI: 0.85–0.89). There was evidence of marked heterogeneity among studies ($p < 0.001$). • Odds ratios of 0.50 (95% CI: 0.26–0.94) for infants, 0.90 (95% CI: 0.87–0.92) for young children, 0.66 (95% CI: 0.60–0.72) for older children, and 0.80 (95% CI: 0.71–0.91) for adults were observed (test for trend, $P = .85$, adjusted for study size; $P = .99$ with the exclusion of infants). • In 6 studies, it was possible to examine the effect of adjustment for the following potentially important confounders: socioeconomic status, parental BMI, and current maternal smoking or maternal smoking in early life. The pooled odds ratio in these studies was reduced from 0.86 (95% CI: 0.81–0.91) before adjustment to 0.93 (95% CI: 0.88–0.99) after combined adjustment. • In 14 studies with information on breastfeeding duration, the effect of breastfeeding over formula feeding was greater among subjects breastfed for ≥ 2 months (OR: 0.81; 95% CI: 0.77–0.84), compared with those breastfed for any duration (OR: 0.89; 95% CI: 0.86–0.91) in the same studies. • Studies that did not provide odds ratios were much less likely to report that breastfeeding was associated with a reduced risk of obesity, compared with studies that did provide odds ratios (1 of 35 studies and 18 of 29 studies, respectively; $P < .001$).	
Quality of the systematic review		B	
Author's interpretations of the results		“The association between breastfeeding and obesity could reflect selective reporting and/or publication. Our results indicated selective reporting of odds ratios by studies that showed a relationship between breastfeeding and reduced risk of obesity. However, because the studies that did not present odds ratios were on average much smaller than those that presented data, their inclusion had a minimal impact on effect estimates.”	
Comments / Limitations		Why not use random-effect model?	

Author, yr: Owen, 2005

Topic of the systematic review*/MA:

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Yes
2	Types of exposure or intervention used	Yes
3	Types of study designs used	Yes
4	Study population	Yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Yes
6	Effort to include all available studies (contact with authors)	Yes
7	Databases and registries searched	Yes
8	Method of addressing published articles other than English	No
9	Method of handling abstracts and unpublished studies	No
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Yes
11	Assessment of confounding	Partially
12	Assessment of quality	Yes
13	Assessment of heterogeneity	Yes
14	Description of statistical methods sufficient to replicate	Yes
15	Provision of appropriate tables and graphs	Yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Yes
17	Tables giving descriptive information of each study included	Yes
18	Results of sensitivity testing (subgroup analysis)	Yes
19	Indication of statistical uncertainty of findings	Yes
20	Quantitative assessment of bias (eg. publication bias)	Yes
	Overall quality	B

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Yes
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	Yes
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Yes
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Partially
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Yes
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Yes
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Yes
9	Were conclusions justified by the reported/collected data and analysis?	Yes

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Sudden Infant Death Syndrome

Hauck 2003

McVea 2000 SRMA*

Mitchell 1997

Schellscheidt 1997

Venneman 2005

* Systematic Review/Meta-Analysis

Hauck, 2003 [UI# 12728140]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 89 days for cases and 85 days for controls Mean BW (range): 2813 g cases and 2915g for controls % Male: nd Race: 75% black; 13.1% Hispanic white; 11.9% non-Hispanic white Enrolled/Evaluate: 260 SIDS and 260 matched controls Location: community Sites: Single/Multi Population based Funding: Gov	Case-control study Follow-up NA	Cases: Included Chicago resident infants whose death between Nov 1993 and April 1996 was determined by the office of medical examiner of Cook County, Illinois to be caused by SIDS Two weeks after caregivers were asked 235 questions Controls: One living control infant was matched to each case infant on maternal race/ethnicity, age at death/interview; and birth weight. They were randomly selected in groups of 20 for white infants and in groups of 40 for Hispanic and black infants	Breastfeeding current yes and ever yes	Breastfeeding current no and ever no

Hauck, 2003 [UI# 12728140]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
SIDS: The sudden death of an infant under 1 yr of age, which remains unexplained after a thorough case investigation, including performance of a complete autopsy, examination of the death scene, and review of the clinical history.	Univariate and multivariate models Adjusted for maternal age, marital status, education, and index of prenatal care	Breast feeding No was reference Breast feeding ever No (cases v controls)/Yes (cases v controls) N=205 v 130 / 55 v 130 OR=0.2 (0.1-0.3) Adj OR=0.4 (0.2-0.7) Breastfeeding current No (cases v controls)/Yes (cases v controls) N=243 v 199 / 17 v 61 OR=0.2 (0.1-0.4) Adj OR=0.3 (0.2-0.7)	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design		x	
			Confounder	x		
			Blinding			
			Data collection			
			Withdraw and dropout			
			Analyses			
			Intervention integrity		x	

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI]	McVea 2000 [11138219]	SIDS
Topic	The role of breastfeeding in sudden infant death syndrome (SIDS)	
Key questions addressed	Question 1 – Infant mortality	
Dates of literature search –	1966-1997	
Databases searched	MEDLINE-all English language literature search Dissertation Abstracts Online Bibliography of all relevant articles	
Countries where primary studies conducted (Developed v developing)	Developed (Regions - North America, Europe, and Australia)	
Study eligibility / inclusion criteria	Minimal definition of SIDS was met (sudden unexplained death of an infant or young child) Original data presented that allowed calculation of an odds ratio for bottle feeding	
Study design [No. Of studies]	Total = 23 Case control [18] Nested case control [4] Cohort [1]	
No. of subjects	Cases (n=99) Controls (n=47 413)	
Study population (definition in included studies)	Children (Age range birth to 2 years when outcome occurred – 17 studies; 8 studies nd on age range)	
Intervention (definition in included studies)	Breast feeding (No standard method of measuring breast feeding was used in the studies - breastfed at birth, breastfed at the time of death, “breast fed”, “partially breast fed”)	
Comparator (definition in included studies)	Bottle fed (nd)	
Outcomes (definition in included studies)	SIDS (autopsy proven in majority – 14 studies, death certificates or coroner’s reports – 4 studies, unclear – 6 studies)	
Heterogeneity	Not explored although studies were heterogeneous with respect to a variety of factors	
Quality of Individual studies	Yes. Quality scores of studies ranged from 11 to 36 points (on a 42 point max score) Better quality studies reported slightly higher risks of bottle feeding	
Publication bias	Yes. “The shape of the funnel graph endorsed the absence of publication bias of only positive studies”	
Statistical Analysis	Random Effects model Used unadjusted odds ratio for the meta-analysis Separate meta-analysis was performed using only those studies with “good” quality scores and after 1988 Dose-response relationship explored	
Results	The pooled OR for the 23 studies using random effects model resulted in an OR = 2.11 (95% CI 1.66-2.68) i.e. the overall risk of SIDS was twice as great for bottle-fed infants compared to breastfed infants. The pooled OR from the higher quality studies also demonstrated a two fold increase in risk among bottle fed infants OR = 2.24 The pooled OR from studies after 1988 OR = 2.32 Confounders: Individual studies adjusted for potential confounders; 6 studies reported adjusted OR; 4 studies reported no protective effect of breastfeeding, while 2 reported adj OR that remained significant. Dose Response relationship: 4 out of 9 studies showed a dose response trend with the risk of SIDS increasing with increasing formula feeding. None of the studies had sufficient power to demonstrate a statistically significant difference between partial v no breastfeeding.	
Quality of the systematic review	C (see limitations)	

Author, Year[UI]	McVea 2000 [11138219]	SIDS
Comments / Limitations	Misclassification of SIDS and breastfeeding Unadjusted odds ratio for pooled analysis Heterogeneity not explored Unable to combine studies with adjusted OR in the pooled analysis because of differences in case matching	

Author, yr: McVea, 2000

Topic of the systematic review*/MA: The role of breastfeeding in SIDS

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Y
2	Types of exposure or intervention used	Y
3	Types of study designs used	Y
4	Study population	Y
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Y
6	Effort to include all available studies (contact with authors)	N
7	Databases and registries searched	Y
8	Method of addressing published articles other than English	N
9	Method of handling abstracts and unpublished studies	Y
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Y
11	Assessment of confounding	N
12	Assessment of quality	Y
13	Assessment of heterogeneity	N
14	Description of statistical methods sufficient to replicate	Y
15	Provision of appropriate tables and graphs	Y
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Y
17	Tables giving descriptive information of each study included	Y
18	Results of sensitivity testing (subgroup analysis)	N
19	Indication of statistical uncertainty of findings	Y
20	Quantitative assessment of bias (eg. publication bias)	N
	Overall Quality	C

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	partially
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	partially
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	no
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	yes
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	yes
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	yes
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	no
8	Were results reported accurately (eg, no discrepancies between text and tables)?	yes
9	Were conclusions justified by the reported/collected data and analysis?	partially

Mitchell 1997 [UI# 9346984]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 2.6-9.1 wk for cases; 2.4-9.2 wk for controls Mean BW (range): <2500 to 3500+ % Male: 60% cases; 52% controls Race: Maori 56% cases and 22% controls Non Maori 44% cases and 78% controls Enrolled/Evaluate: 232 cases and 922 controls Location: community Sites: Single/Multi Population based Funding: Gov	Case-cohort study design 2 yr cohort; follow-up duration NA	Cases: All deaths registered by the New Zealand Health Information Service as attributable to SIDS in the postneonatal age group (dying after 28 completed days and within the first year of life) form the cases of this study. 98% of deaths classified as SIDS have had a autopsy. Controls: Randomly selected sample from a prospectively collected cohort on all births in the 2 yr period. The method of sampling was 1) a date of birth was randomly selected from all the days in the study period; 2) an obstetric hospital was randomly chosen in proportion to the number of births; 3) in the obstetric hospital with multiple births on nominated date of birth random numbers were used to select a particular infant from among those born on that day; 4) a direction variable which indicates to either go forward or back in looking for a birth in the situation where the hospital did not have one on the nominated day was also randomly chosen	Three measures of breastfeeding reference yes Exclusive breastfeeding at discharge from the obstetric hospital; any breastfeeding at initial contact; and any breast feeding at 2 months	No for the three measures of the breastfeeding

Mitchell 1997 [UI# 9346984]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
SIDS: dying after 28 completed days and within the first year of life	Univariate and multivariate analysis Controlled for: maternal age, marital status, age mother left school, previous number of pregnancies, infant's sex, ethnicity of infant, birthweight, sleep position, breastfeeding and bed sharing/maternal smoking combinations.	The three measures of breastfeeding were not statistically associated with a statistically significant reduction in the risk of SIDS	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design	x		
			Confounder	x		
			Blinding			
			Data collection			
			Withdraw and dropout			
			Analyses			
Comments	The controls were randomly selected from a prospective cohort; however the analysis was conducted as a case-control study. Overall quality = B		Intervention integrity			x

Schellscheidt J 1997 [UI# 9266202]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 8 days to one yr Mean BW (range): nd % Male: nd Race: nd Enrolled/Evaluate: 59 Location: community Sites: Single/Multi Population based study Funding: nd	Case-control study Follow-up Not applicable Incidence of SIDS during the 2 yrs	For cases: No positive autopsy findings; minimal findings without any relationship to death; findings possibly related to death (but not explaining it) and autopsy findings explaining death. For controls: Two controls were selected for each case of the same age ± 4 weeks and sex randomly by the pediatrician or a general practitioner	Breastfeeding in weeks >12 wks (ref)	Breastfeeding in weeks 1-6 wks 7-12 wks None

Schellscheidt J 1997 [UI# 9266202]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Same as eligibility criteria	Variables were entered as dichotomous indicators Analyses were univariate. No confounders were adjusted	Breast feeding cases versus controls None (N=29 vs 27) OR=7.7 (2.7, 22.3) 1-6 wk (N=12 vs 40) OR=2.1 (0.7, 6.7) 7-12 wk (N=10 vs 39) OR=1.8 (0.6, 6.0) >12 wk (N=7 vs 50) REF	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design	x		
			Confounder			x
			Blinding		x	
			Data collection		x	
			Withdraw and dropout			
			Analyses			x
			Intervention integrity			

Vennemann, 2005 [UI# 16188764]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): nd Mean BW (range): nd % Male: cases 60.4% Controls 60.3% Race: nd Enrolled/Evaluate: SIDS cases 404/333 Location: population based Sites: Single/Multi 18 centers involved Funding: nd	Case-control study design Duration SIDS during 3 yr period Follow-up duration NA	Cases: All reported cases of sudden and unexpected deaths in the first year of life after the first 7 d and all cases were autopsied. Controls: Three controls with the best age matching were chosen	Breastfeeding >2 wk yes	Breastfeeding >2wk No

Vennemann, 2005 [UI# 16188764]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Sudden and unexpected deaths in the first year of life after the first 7 d and all cases were autopsied.	Univariate and multivariate analysis were done using conditional logistic regression. The multivariate model includes all variables which were found significant at the 5% level in the univariate analysis except gestational age, as this was closely related to birthweight	Breastfeeding >2wk Yes N=165/827 (cases/controls) No N=168/171 Univariate OR=5.36 (3.97-7.23) Multivariate OR=1.71 (1.06-2.77)	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design	x		
			Confounder		x	
			Blinding		x	
			Data collection		x	
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity		x	
			Overall = B			

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Type 1 DM

EURODIAS 2002
Gerstein 1994 SRMA*
Jones 1998
McKinney 1999
Meloni 1997
Norris 1996 SRMA*
Tai 1998
Visalli 2003

* Systematic Review/Meta-Analysis

EURODIAS, 2002 [12351473]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Age at Dx of IDDM: ND Mean GA (range): ND Mean BW (range): ND % Male: ND Race: ND Enrolled/Evaluate: 683/610 for cases; 2167/1616 for controls Location: Europe Sites: Multi Funding: Government	Case-control study	Cases: cases were identified from a population-based register of childhood-onset diabetes operating in a sub-study of EURODIAB (5 participating centers) Controls: a population-based sample of control children, matched to the patients in age distribution, was obtained in each center using sources that depended on local circumstances as previously described (including schools, population register, polyclinics, and general practitioner register)	Parents were interviewed or filled in the questionnaire regarding information about infant feeding (duration of breast-feeding, age at introduction of formula feeding, dairy milk, food containing fruit, vegetable, fish, meat, and egg). Breast-feeding of any duration (or ever breast-feeding)	Never breast-feeding

EURODIAS, 2002 [12351473]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Pooled ORs of IDDM for exposures obtained from each center	Mantel Haenszel approach was used to pool odds ratio (ORs) for exposures (e.g. breastfeeding, early introduction of solid food) obtained from each center, to test the significance of the combined OR, and to test for heterogeneity in the ORs between centers. To adjust for potential confounders, logistic regression analysis was used with terms included in the model to represent centers. Adjustments included height SDS, weight SDS, maternal age at delivery, neonatal jaundice, neonatal respiratory infection, vitamin D supplementation, and asthma.	Breast-feeding of any duration was associated with a reduction in risk, with a pooled OR of 0.75 (95% CI 0.58-0.96) and no evidence of heterogeneity between centers. The introduction of cow's milk before 3 months of age (OR 1.15, 95%CI 0.74-1.81), cow's milk or formula (OR 1.01, 95%CI 0.81-1.25), or solid foods (OR 0.74, 95%CI 0.57-0.95) was not associated with any significant elevation in risk. Indeed the finding for solid foods suggested a reduced risk, although there was significant heterogeneity between centers (p<0.001). The OR associated with breastfeeding in multivariate analysis remained significant (OR=0.59, 95%CI 0.35-0.97)	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder	x		
			Blinding			x
			Data collection		x	
			Withdraw and dropout		x	
			Analyses	x		
			Intervention integrity	N/A		
			Overall: B Long-term recalls			

Q1: Insulin-Dependent Diabetes Mellitus

Author, Year[UI]	Gerstein, 1994 [8112184]	Type 1 Diabetes Mellitus
Literature search (Dates)	Medline (no data); Other databases searched? (no); unpublished data used? (yes)	
Countries where primary studies conducted	Developed countries only: Norway, Denmark, Italy, US, Australia, Canada, Sweden, Finland, UK	
Study eligibility / inclusion criteria	Articles or letters to the editor were eligible for inclusion if they reported original research dealing with both type I diabetes and cow's milk exposure or avoidance and were published in peer-reviewed journals. Articles were excluded if they exclusively used surrogate markers for either type I diabetes or cow's milk exposure.	
Study design [No. Of studies]	Ecological and time-series studies [3]; Case-control studies [13]; Cohort study [1]; case series [1]	
No. of subjects	Data available for case-control studies only: Cases: 3,708 Controls: 20,340	
Study population (definition in included studies)	• Ecological and time-series studies in which the prevalence of type I DM was compared with the rate of breast-feeding or cow's milk consumption in different populations or over a specified period of time. • Case-control studies in which the neonatal feeding histories of patients with type I DM and individually matched non-DM control subjects were compared. Studies that used population breast-feeding registries as control data also were included, because the low prevalence of type I DM in the general population ensures that population data is representative of the breast-feeding habits of non-DM individuals. • There was a letter to the editor reporting an analysis of two cohort studies and an article that described the neonatal feeding history of a series of patients with DM. • There were no controlled trials of exposure to cow's milk.	
Intervention/Exposure (definition in included studies)	Cow's milk exposure (and therefore non-exclusive breastfeeding)	
Comparator (definition in included studies)	Cow's milk avoidance (and therefore breastfeeding)	
Outcomes (definition in included studies)	Type I DM. Surrogate markers of type I DM (such as the presence of islet cell antibodies) were excluded.	
Heterogeneity assessments	Q statistics and chi-square distribution	
Quality assessments	Six methodological criteria were assessed for case-control studies only. Each criterion was marked individually and listed in a summary table of all studies. There was no summary grading.	
Publication bias assessments	Not done	
Statistical Analysis or meta-analytic methods	Meta-analyses of ORs were done for case-control studies. Both adjusted (preferred) and unadjusted ORs were used. Fixed-effect model.	

Author, Year[UI]	Gerstein, 1994 [8112184]	Type 1 Diabetes Mellitus
Results	- Time-series and ecological studies showed an association (geographically and temporally) between rate/prevalence of breastfeeding and the incidence of type I DM. - Results from the 13 case-control studies were mixed. Some of these studies demonstrated that patients with type I DM were more likely to have a history of neonatal cow's milk exposure or a short/negative history of breast-feeding than non-DM control subjects; other studies failed to demonstrate such a relationship. When these results were meta-analyzed, the OR for type I DM in patients exposed to <3 months of breast-feeding was 1.38 (95% CI 1.22-1.53; p=0.11 for homogeneity). Similarly, the overall OR for cow's milk exposure before 3-4 months of age in patients with type I DM was 1.57 (95% CI 1.19-2.07; p=0.10 for homogeneity). - An analysis of two cohorts of children born in the UK in 1958 and 1970 followed for 16 and 10 years, respectively, failed to show any association between breastfeeding for <1 month and type I DM.	
Quality of the systematic review	B	
Author's interpretations of the results	The results of this review are consistent with the hypothesis that avoidance of cow's milk products during the first few months of life may reduce the risk of type I DM. A feeding intervention trial in which susceptible newborns would be randomized to receive formula with or without cow's milk and followed for the development of DM would most clearly resolve the issue.	
Comments / Limitations	This systematic review primarily aimed to examine the relationship between early cow's milk exposure and type I DM. However, exclusive breastfeeding is a way of avoidance of cow's milk exposure; then therefore this review also examined the association between breastfeeding and type I DM. The exclusivity of breastfeeding was not addressed in the meta-analyses. Combining both adjusted and crude OR.	

Author, yr: Gerstein, 1994

Topic of the systematic review*/MA: Type 1 DM

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Yes
2	Types of exposure or intervention used	Yes
3	Types of study designs used	Yes
4	Study population	Yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	No
6	Effort to include all available studies (contact with authors)	No
7	Databases and registries searched	Yes
8	Method of addressing published articles other than English	No
9	Method of handling abstracts and unpublished studies	No
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Yes
11	Assessment of confounding	No
12	Assessment of quality	Yes
13	Assessment of heterogeneity	Yes
14	Description of statistical methods sufficient to replicate	Yes
15	Provision of appropriate tables and graphs	Yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Yes
17	Tables giving descriptive information of each study included	Yes
18	Results of sensitivity testing (subgroup analysis)	Yes
19	Indication of statistical uncertainty of findings	Yes

		Reporting yes/no
20	Quantitative assessment of bias (eg. publication bias)	Yes
	Overall quality	B

*When grading a systematic review, please skip item #14, 18-20
Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Partially
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	No
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Yes
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Partially
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Partially
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Yes
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Yes
9	Were conclusions justified by the reported/collected data and analysis?	Yes

Jones, 1998 [9698133]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Age at Dx of IDDM: ND Mean GA (range): ND Mean BW (range): ND % Male: 53 Race: ND Enrolled/Evaluate: <76/60 for cases Location: UK Sites: Multi Funding: Government	Case-control study	Cases: Oxford Record Linkage Study (ORLS) records of hospital admissions for discharge diagnoses coded to diabetes mellitus were searched. Cases were identified as children who were diagnosed with diabetes 1965-1987, and who had been born during 1965-1986 (N=315). 98% of cases born 1970-1986 were matched with 5 or more controls each. Controls: for cases born 1970-1986, 8 controls from all livebirths in the ORLS areas were randomly selected. Controls were individually matched to cases on sex, year and hospital of delivery, or if domiciliary to another domiciliary delivery. Exclusion: children with cystic fibrosis or major congenital anomalies, or who were part of twin or higher order deliveries. Breastfeeding data were routinely collected only from 1976 so this result is based upon 60 cases and 458 controls born 1976-1986, of which 7 controls had missing data.	Breastfeeding at discharge	Not breastfeeding at discharge

Jones, 1998 [9698133]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Relative risk of IDDM	Relative risks were calculated by estimating rate ratios using conditional logistic regression for matched case-control studies.	There was a small, but not significant, raised risk of diabetes with not breastfeeding at discharge (RR=1.33, 95% CI: 0.76-2.31).	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design		x	
			Confounder			x
			Blinding		x	
			Data collection			x
			Withdraw and dropout			x
			Analyses			x
			Intervention integrity	N/A		
			Overall: C Poor adjustments for confounding; Poor definition of BF exposure			

McKinney, 1999 [10372244]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Age at Dx of IDDM: 0-15 Mean GA (range): 11% ≤37 wks Mean BW (range): 7% <2500g, 58% 2500-3500g; 35% ≥3500g % Male: 55 Race: 95% White, 3.5% India or Pakistan Enrolled/Evaluate: 202/196 for cases/354/325 for controls Location: UK Sites: Single Funding: Private	Case-control study	Cases: Incident cases included those diagnosed with type 1 DM over a 2-year period (1993-1994) and taken from the population-based Yorkshire Childhood Diabetes Register (YCDR). The register ascertained children from 3 independent sources and was estimated to be 97% complete. Controls: 2 control subjects per case, matched by sex and age, were randomly selected from the primary care registrations of the Family Health Service Authority (FHSA) of the matched case. Exclusion: 6 cases and 29 control subjects had no matching control subjects or cases and were lost to the matched analysis	Mothers were interviewed. Participation rates for the interview study were 93.6% and 81.9% for case and control mothers, respectively. Initial exclusive breast feeding (yes): mothers were asked how the infant was first fed	Initial exclusive breast feeding (no)

McKinney, 1999 [10372244]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
OR of IDDM	ORs with 95% CI and p values were calculated using conditional logistic regression with a single term in the model for univariate estimates. Variables significantly associated with disease obtained from the analysis of individual variables were modeled together in a multivariable analysis. Variables significant at the 5% level in a univariate analysis in our study were mother's age (<25, 25-35, >35), type 1 diabetes mothers, preeclampsia, delivery by cesarean section, neonatal illnesses, and breast feeding. These potential explanatory variables were modeled together in a multivariable analysis.	Factors in the model	Cases	Controls	OR (95% CI)	A: strong, B: moderate, C: weak	A	B	C
		n	196	325					
		Mother's age							
		25-35 years	123	182	1.69 (1.11-2.60)	Selection	x		
		>35 years			2.07 (0.97-4.43)	Study design		x	
		Mother with IDDM	4	0	--	Confounder		x	
		Preeclampsia	44	49	1.60 1.01-2.54)	Blinding			x
		Cesarean delivery	34	35	1.45 (0.82-2.55	Data collection			x
		Neonatal illnesses ≥1	61	35	1.55 (1.00-2.42)	Withdraw and dropout	x		
		Initial exclusive breast feeding	73	151	0.60 (0.41-0.89)	Analyses	x		
		Results show all variables retained their significance in the multivariate model. The addition of neonatal illnesses and breast feeding significantly improved the model fit (likelihood ratio test, chi squared 9.96, P = 0.007).				Intervention integrity	N/A		
		Note: Crude OR = (73*174)/(151*123)=0.68, p=0.04				Overall: B No adj. for SES			

Meloni, 1997 [9051384]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Age at Dx of IDDM: 6 (1-15) yr Mean GA (range): ND Mean BW (range): ND % Male: 53 Race: ND Enrolled/Evaluate: 115/100 for cases Location: Italy Sites: Single Funding: ND	Case-control study	Cases: All diabetic subjects followed up at the university hospital who met the inclusion criteria: 1: age < 17 years at diagnosis of IDDM, 2) diagnosis of IDDM was achieved between 1983 and 1994, and 3) patient's mother was currently living, which was considered necessary to obtain accurate infant diet information. According to these criteria, 100 IDDM patients (of the 115 eligible) were included. Controls: children admitted at the same university hospital, matched each diabetic case for sex and age. None of the control subjects had a family history of IDDM.	Long-term maternal recall of the duration of complete or partial breast-feeding during infant's 1 st year of life.	Long-term maternal recall of the age at which dietary products containing cow's milk were introduced into the diet during infant's 1 st year of life

Meloni, 1997 [9051384]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Risk of IDDM, compared children who had been breast-fed to those who had not been breast fed	The OR of diabetes according to breast-feeding, the duration of breast-feeding, the age at introduction of dietary products containing cow's milk, and the age at introduction of solid food were estimated by means of conditional logistic regression, where the strata were defined by single year of age at interview and sex. Also included in the models were terms for mother's education and number of siblings.		IDDM cases	Controls	OR (95% CI)	A: strong, B: moderate, C: weak	A	B	C
Risk of IDDM, compared children who had been breast-fed for 1-2 months, 3-5 months, or >6 months to those who had not been breast-fed.		BF							
		Yes	84	70	1*			x	
		No	16	30	0.41 (0.19-0.91)			x	
		Duration of BF (mo)						x	
		>6	25	19	1*				x
		3-5	41	25	1.18 (0.52-2.68)		x		
		1-2	18	26	0.48 (0.19-1.24)		x		
		0	16	30	0.36 (0.14-0.94)				
		Continuous coefficient**			1.10 (0.99-1.22)		x		
		*reference category					Intervention integrity		
		**continuous per 1 month increase					N/A		
		Overall: B							
Limitations: Long-term maternal recalls. Hospital controls									

Q1: Insulin-Dependent Diabetes Mellitus

Author, Year[UI]	Norris, 1996 [8664407]	Type 1 Diabetes Mellitus
Literature search (Dates)	Medline (1966-1994); Other databases searched? (no); unpublished data used? (yes)	
Countries where primary studies conducted	Developed countries only: Norway, Denmark, Italy, US, Australia, Canada, Sweden, Finland, British Isles, Ireland, UK	
Study eligibility / inclusion criteria	Inclusion: all epidemiologic studies examining infant diet and IDDM risk Exclusion: studies with insufficient data for meta-analyses	
Study design [No. Of studies]	Case-control studies [17]	
No. of subjects	Cases: 4,656 Controls: 16,383	
Study population (definition in included studies)	Cases: IDDM patients identified from diabetes clinics, school surveys, registries, physician surveys, or pediatric departments. Controls: non-IDDM controls selected from siblings, national population data, friends, classmates, physician referred samples, national population registry, random sample of household, child health system, or schools and day care	
Intervention/Exposure (definition in included studies)	A=breastfeeding status (ever/never) B=total breastfeeding duration	
Comparator (definition in included studies)	C=exposure to breast-milk substitutes D=exposure to cow's milk-based substitutes	
Outcomes (definition in included studies)	Odds ratios for the risk of IDDM	
Heterogeneity assessments	"The addition to the model of a variable for each study in the analyses resulted in an improvement in the overall fit of the model (as determined by the log-likelihood statistic) compared with the model containing only the exposure variable, indicating heterogeneity of effect across studies. We also evaluated study heterogeneity by adding to the model product terms obtained by multiplying the exposure variable by each study variable."	
Quality assessments	Sensitivity analyses on various study characteristics	
Publication bias assessments	Including unpublished data	
Statistical Analysis or meta-analytic methods	Stratified unconditional logistic regression was used to estimate OR and 95%CI. This method weights each study by its sample size. Variables for the individual studies and for the exposure of interest were included in the model. Both matched and unmatched ORs were used.	
Results	- The summary odds ratio (OR) of the 18 studies that examined the association between never being breast-fed and IDDM was 1.13 (95% CI = 1.04-1.23). - Meta-analyses of summary ORs were done in the 14 studies that examined IDDM risk by months of breastfeeding duration. The duration categories in the analyses were cumulative, rather than mutually exclusive. The summary OR for IDDM in subjects who were breast-fed for less than 3 months compared with those who were breast-fed for at least 3 months was 1.23 (95%CI = 1.12-1.35). - When comparing the ORs by whether or not the studies relied on long-term recall to assess infant diet, studies using existing infant records to determine breastfeeding initiation and duration failed to show the associations reported in the studies relying on long-term recall for their exposure data.	
Quality of the systematic review	B	
Author's interpretations of the results	Our meta-analysis showed that the increased risk of IDDM associated with any of the infant diet exposures is small Interpretation of weak associations (that is, generally an OR<2) can be problematic, since weak associations can more readily be explained by bias.	

Author, Year[UI]	Norris, 1996 [8664407]	Type 1 Diabetes Mellitus
Comments / Limitations	Very thoughtful analyses but the analysis cannot be replicated due to lack of individual study data. Only crude ORs were used in the meta-analyses due to limitation of meta-analytic technique. Conclusions were well justified.	

Author, yr: Norris, 1996

Topic of the systematic review*/MA: Type 1 DM

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	yes
2	Types of exposure or intervention used	yes
3	Types of study designs used	yes
4	Study population	yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	no
6	Effort to include all available studies (contact with authors)	yes
7	Databases and registries searched	yes
8	Method of addressing published articles other than English	no
9	Method of handling abstracts and unpublished studies	yes
	Reporting of Methods	
10	Rationale for selection and/or coding of data	yes
11	Assessment of confounding	yes/no
12	Assessment of quality	yes
13	Assessment of heterogeneity	yes
14	Description of statistical methods sufficient to replicate	yes
15	Provision of appropriate tables and graphs	yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	no
17	Tables giving descriptive information of each study included	yes
18	Results of sensitivity testing (subgroup analysis)	yes
19	Indication of statistical uncertainty of findings	yes
20	Quantitative assessment of bias (eg. publication bias)	yes
	Overall quality	B

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Partially
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	Yes
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Yes
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Partially
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Yes
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Yes
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Yes
9	Were conclusions justified by the reported/collected data and analysis?	Yes

Tai, 1998 [9925351]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Age at Dx of IDDM: 8.3±3.3 SD Mean GA (range): 1% ≤36 wk; 99% >36 wk Mean BW (range): 21% <3 kg, 69% 3-3.9 kg; 10% ≥4 kg % Male: 37 Race: Asian Enrolled/Evaluate: 117/117 for cases; 193/193 for controls Location: Taipei, R.O.C. Sites: Single Funding: Government	Case-control study	Cases: From 1984 to 1993, there were a total of 161 registered type 1 diabetes patients who were under the age of 30, born in Taipei City and continuously living in the same city. Among these, 119 (73%) subjects agreed to participate in the study. Their overnight fasting serum C-peptide concentration was < 0.12 nmol/l and/or an increment in C-peptide from basal value after 1mg glucagons intravenous injection. Controls: The control group was selected from normal classmates or colleagues of the type 1 diabetes subjects. They were frequency-mated for age, sex, and parental and individual educational levels. All controls had a plasma glucose concentration, 2 h after 1.75 g/kg glucose loading, <140 mg/dl, fasting C-peptide value >0.16 nmol/l and were negative for islet cell antibody.	In the majority of the study subjects (79% cases and 75% controls), their mothers were interviewed by 2 trained nurses. In the remaining study subjects, the questionnaire interviews were answered by the subjects themselves or by their fathers with the help of their mothers. Feeding style during infancy: breast-feeding Duration of breast-feeding: <6, ≥6 months	Never breast-fed (totally on cow's milk)

Tai, 1998 [9925351]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments																																								
OR of developing IDDM	Age and sex were adjusted using logistic regression analysis. The age-sex-adjusted ORs were estimated from the regression coefficient and its S.E. for the risk factor. Multiple logistic regression analysis was then carried out by including risk factors which were statistically significant in univariate analyses as well as those reported in other studies in the model.	I. Age-sex-adjusted ORs of developing IDDM (univariate):			<table><tr><td>A: strong, B: moderate, C: weak</td><td>A</td><td>B</td><td>C</td></tr><tr><td>Selection</td><td></td><td>x</td><td></td></tr><tr><td>Study design</td><td></td><td>x</td><td></td></tr><tr><td>Confounder</td><td>x</td><td></td><td></td></tr><tr><td>Blinding</td><td></td><td></td><td>x</td></tr><tr><td>Data collection</td><td></td><td></td><td>x</td></tr><tr><td>Withdraw and dropout</td><td></td><td>x</td><td></td></tr><tr><td>Analyses</td><td></td><td>x</td><td></td></tr><tr><td>Intervention integrity</td><td colspan="3">N/A</td></tr></table> Overall: B Inconsistent methods for BF exposure ascertainment			A: strong, B: moderate, C: weak	A	B	C	Selection		x		Study design		x		Confounder	x			Blinding			x	Data collection			x	Withdraw and dropout		x		Analyses		x		Intervention integrity	N/A		
		A: strong, B: moderate, C: weak	A	B				C																																			
		Selection		x																																							
		Study design		x																																							
		Confounder	x																																								
		Blinding						x																																			
		Data collection						x																																			
		Withdraw and dropout		x																																							
		Analyses		x																																							
		Intervention integrity	N/A																																								
		<table><tr><td>Risk factors</td><td>Cases N</td><td>Controls N</td><td>OR (95% CI)</td></tr><tr><td colspan="4">Feeding style during infancy</td></tr><tr><td>Breast-feeding</td><td>46</td><td>82</td><td>1.00</td></tr><tr><td>Cow's milk</td><td>70</td><td>109</td><td>1.43 (0.83-2.47)</td></tr><tr><td colspan="4">Duration of BF (months)</td></tr><tr><td>Never</td><td>70</td><td>109</td><td>1.00</td></tr><tr><td><6</td><td>32</td><td>52</td><td>0.82 (0.47-1.42)</td></tr><tr><td>≥6</td><td>10</td><td>25</td><td>0.39 (0.16-0.94)</td></tr></table>			Risk factors	Cases N	Controls N	OR (95% CI)	Feeding style during infancy				Breast-feeding	46	82	1.00	Cow's milk	70	109	1.43 (0.83-2.47)	Duration of BF (months)				Never	70	109	1.00	<6	32	52	0.82 (0.47-1.42)	≥6	10	25	0.39 (0.16-0.94)							
		Risk factors	Cases N	Controls N	OR (95% CI)																																						
		Feeding style during infancy																																									
		Breast-feeding	46	82	1.00																																						
		Cow's milk	70	109	1.43 (0.83-2.47)																																						
Duration of BF (months)																																											
Never	70	109	1.00																																								
<6	32	52	0.82 (0.47-1.42)																																								
≥6	10	25	0.39 (0.16-0.94)																																								
II. Multiple logistic regression in which the order of pregnancy, paternal age at the conception of the study subjects, gestational age, type of delivery, birth weight, duration of breast-feeding and monthly family income were included in the model.																																											
<table><tr><td>Duration of BF (months)</td><td>OR (95% CI)</td></tr><tr><td>Never</td><td>1.00</td></tr><tr><td><6</td><td>0.84 (0.45-1.59)</td></tr><tr><td>≥6</td><td>0.25 (0.09-0.69)</td></tr></table>			Duration of BF (months)	OR (95% CI)	Never	1.00	<6	0.84 (0.45-1.59)	≥6	0.25 (0.09-0.69)																																	
Duration of BF (months)	OR (95% CI)																																										
Never	1.00																																										
<6	0.84 (0.45-1.59)																																										
≥6	0.25 (0.09-0.69)																																										
The dose-response relationship with type 1 DM for pregnancy order and the reverse pattern for duration of breast-feeding remained statistically significant after adjustment for multiple risk factors.																																											

Visalli, 2003 [12876166]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Age at Dx of IDDM: ND Mean GA (range): ND Mean BW (range): ND % Male: 50 Race: ND Enrolled/Evaluate: 330/150 for cases; 1200/750 for controls Location: Italy Sites: Multi Funding:	Case-control study	Cases: type 1 DM patients selected (330 cases) for this study were diagnosed within the EURODIAS ACE study, and were born between 1977 and 1989. Controls: population based control subjects born in the same period were chosen from primary and secondary school enrolment records in the region, and the selection process was designed in collaboration with EURODIAS Substudy 2 group. Out of 1,100 controls replied to the questionnaires, 750 were randomly selected, matching 1 case with 5 controls by age.	Questionnaires were distributed to all parents of cases and controls. Duration of breast feeding ≥ 3 months End of breast feeding ≥ 3 months	Duration of breast feeding < 3 months End of breast feeding < 3 months

Visalli, 2003 [12876166]

Outcome Definition	Statistical analyses and confounders adjusted	Results			Bias/limitations Comments			
OR of IDDM	Conditional logistic regression analysis was performed to adjust the possible confounders (those significant in univariate analyses).	I. Univariate analysis:			A: strong, B: moderate, C: weak	A	B	C
		Risk factors	OR	95% CI	Selection			x
		Full term birth	1.44	0.81-2.56	Study design		x	
		Duration of BF < 3 mo	2.12	1.38-3.26	Confounder		x	
		Beginning of weaning < 3 mo	1.81	1.06-3.07	Blinding			x
		End of breast feeding < 3 mo	2.03	1.24-3.33	Data collection		x	
		Cow's milk before 23 mo	1.49	0.72-3.11	Withdraw and dropout			x
		II. Multivariate logistic regression analysis for risk factors which have been found to be significant			Analyses	x		
		Risk factors	OR	95% CI	Intervention integrity	N/A		
		Duration of BF < 3 mo	1.74	1.40-2.45	Overall: C			
		Family history of type 1 DM	1.17	1.05-1.76	Only 45% of the eligible cases were able to be contacted.			
		Presence of infectious disease during mother's pregnancy	1.60	1.32-2.20				
		Occurrence of eczema	1.61	1.25-2.66				

Appendix C. Evidence Tables

Part I. Term Infant Outcomes

Infant Type 2 DM

Owen 2006 SRMA*

Taylor 2005 SRMA*

* Systematic Review/Meta-Analysis

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI] Topic	Owen, 2006 Type 2 diabetes
Literature search (Dates)	Medline (1966-November 2004); Other databases searched? (yes); unpublished data used? (no)
Countries where primary studies conducted	Developed countries: UK, US, Canada, Dutch
Study eligibility / inclusion criteria	Exclusion criteria: infant feeding status or measures of diabetes were not recorded. Studies did not compare diabetic outcomes among those formula and breastfed. Studies did not provide the odds of type 2 diabetes, mean differences in blood glucose, or differences in plasma insulin between those breastfed and formula fed. Note: For the purpose of this review, we only focus on those studies reported OR of type 2 diabetes as an outcomes.
Study design [No. Of studies]	Historical cohort study [3]; Cross-sectional study [2]; Prospective cohort [1]; Case-control study [1]
No. of subjects	Total of 76,744 subjects
Study population (definition in included studies)	6 studies conducted in adults and 1 conducted in adolescents
Intervention/Exposure (definition in included studies)	Ever breastfed vs. formula fed in all 7 studies. Feeding status was reported as being exclusive in one of these studies.
Comparator (definition in included studies)	
Outcomes (definition in included studies)	OR (95% CI) of type 2 diabetes
Heterogeneity assessments	Chi-squared test. Meta-regression and sensitivity analysis were used to examine potential heterogeneity across studies (see Quality assessment)
Quality assessments	The effect of study size, year of birth, the method of ascertainment of infant feeding status (whether contemporary or recalled up to 71 y after birth), type of formula fed, study response rate (analyzed as a continuous variable), study design (randomized controlled trial, case-control, or cohort), and whether infants were born pre- or full-term was examined by using meta-regression and sensitivity analysis. Sensitivity analyses were used to examine the effect of adjustment for important confounders and of fasting status.
Publication bias assessments	Funnel plots with Begg and Egger tests
Statistical Analysis or meta-analytic methods	Fixed-effect model.
Results	• Six of the 7 studies related breastfeeding to a lower risk of type 2 diabetes, and there was no evidence of heterogeneity across studies. • Overall, the subjects who were breastfed showed a lower risk of type 2 diabetes than did those who were formula fed (pooled OR: 0.61; 0.44-0.85; p=0.003). • Three studies had information on relevant confounders (birthweight, parental diabetes, socioeconomic status, and individual or maternal body size). However the OR relating breastfeeding and diabetes risk was similar before (0.55; 95% CI: 0.35-0.86; p=0.009) and after (0.55, 95% CI 0.34-0.90; p=0.017) adjustment. • The method for ascertaining feeding exposure was unrelated to the ORs.
Quality of the systematic review	A
Author's interpretations of the results	In the present overview of published studies relating infant feeding and risk of diabetes, the pooled estimate from 6 studies conducted in adults and 1 study conducted in adolescents showed that early breastfeeding was consistently associated with a lower risk of type 2 diabetes in later life compared with those initially formula fed.

Author, Year[UI]	Topic	Owen, 2006	Type 2 diabetes
		Publication bias is an important potential explanation for the consistent associations observed in these published studies. The results provide no evidence of marked publication bias, although the statistical power of formal tests was limited by the small number of studies available for analysis. It is possible that confounding by birthweight and maternal factors could lead to overestimation of the association between breastfeeding and diabetes in later life.	
Comments / Limitations		Pooled different study designs	

Author, yr: Owen, 2006

Topic of the systematic review*/MA: Type 2 DM

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	yes
2	Types of exposure or intervention used	yes
3	Types of study designs used	yes
4	Study population	yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	yes
6	Effort to include all available studies (contact with authors)	no
7	Databases and registries searched	yes
8	Method of addressing published articles other than English	no
9	Method of handling abstracts and unpublished studies	no
	Reporting of Methods	
10	Rationale for selection and/or coding of data	yes
11	Assessment of confounding	yes
12	Assessment of quality	yes
13	Assessment of heterogeneity	yes
14	Description of statistical methods sufficient to replicate	yes
15	Provision of appropriate tables and graphs	yes
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Yes
17	Tables giving descriptive information of each study included	yes
18	Results of sensitivity testing (subgroup analysis)	yes
19	Indication of statistical uncertainty of findings	yes
20	Quantitative assessment of bias (eg. publication bias)	yes
	Overall quality	A

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Yes
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	Partially
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Yes
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Yes
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Yes
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Yes
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Yes
9	Were conclusions justified by the reported/collected data and analysis?	Yes

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI]	Topic	Taylor, 2005 [UI]	Type 2 diabetes
Literature search (Dates)		Medline (1966-2003); Other databases searched? (yes); unpublished data used? (no)	
Countries where primary studies conducted		Developed countries only: US, Canada, Germany	
Study eligibility / inclusion criteria		Search was limited to the English language and to human subjects. All studies involved term singleton infants unless otherwise noted.	
Study design [No. Of studies]		Cohort studies [2]; case-control study [1]	
No. of subjects		1,514	
Study population (definition in included studies)		- Studies of association between breastfeeding and type 2 diabetes - Children of mothers with diabetes	
Intervention/Exposure (definition in included studies)		Many of the studies that were reviewed only reported breastfeeding versus bottle feeding, without more detailed information on exclusivity, frequency, or duration. Information on timing and duration of breastfeeding were included where available.	
Comparator (definition in included studies)			
Outcomes (definition in included studies)		- Impaired glucose tolerance (IGT): fasting glucose < 140 mg/dL and 2-hour glucose > 140 mg/dL (after 1.75 g/kg OGTT) - Type 2 diabetes (NIDDM): 2-hour glucose ≥ 200 mg/dL (after 75 g OGTT) or fasting glucose ≥ 126 mg/dL	
Heterogeneity assessments		N/A	
Quality assessments		N/A	
Publication bias assessments		N/A	
Statistical Analysis or meta-analytic methods		N/A	
Results		Association between breastfeeding and type 2 DM: - Consistently in 2 studies (1 retrospective cohort; 1 case-control study), exclusive breastfeeding is associated with lower risk of diabetes. Children of Mothers with DM: - One retrospective cohort study found that DM in the next generation was less common among breastfed children than among bottle-fed children, of both mothers with DM and without DM.	
Quality of the systematic review		C	
Author's interpretations of the results		Based on the literature reviewed, breastfeeding should be strongly encouraged for all women, with or without diabetes.	
Comments / Limitations		No synthesis of results or appraisals of study quality. Unclear how the conclusions were reached.	

Author, yr: Taylor, 2005

Topic of the systematic review*/MA: Type 2 DM

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	yes
2	Types of exposure or intervention used	yes
3	Types of study designs used	yes
4	Study population	yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	yes
6	Effort to include all available studies (contact with authors)	no
7	Databases and registries searched	yes
8	Method of addressing published articles other than English	no
9	Method of handling abstracts and unpublished studies	no
	Reporting of Methods	
10	Rationale for selection and/or coding of data	yes
11	Assessment of confounding	partially
12	Assessment of quality	no
13	Assessment of heterogeneity	N/A
14	Description of statistical methods sufficient to replicate	N/A
15	Provision of appropriate tables and graphs	no
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	no
17	Tables giving descriptive information of each study included	no
18	Results of sensitivity testing (subgroup analysis)	N/A
19	Indication of statistical uncertainty of findings	yes
20	Quantitative assessment of bias (eg. publication bias)	no
	Overall quality	C

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Partially
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	No
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	No
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Partially
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	No
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Partially
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Unclear (no tables)
9	Were conclusions justified by the reported/collected data and analysis?	No

Appendix C. Evidence Tables

Part II. Preterm Infant Outcomes

Cognitive Premature Infants

Bier 2002
Eidelman 2004
Elgen 2002
Feldman 2003
Horwood 2001
O'Connor 2003
Pinelli 2003
Smith 2003
Vohr 2006

Bier 2002 [2003020034]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range) breast milk group: 28.6 wk (23-34) Mean BW (range): 1174 g (600-1965) % Male: 45% Race: Enrolled/Evaluate: 29 in human milk group; 10 in formula group Location: US Sites: Multi Funding:	Non-randomized comparison of breast milk versus formula fed preterm infants (convenience sample)	Excluded infants with mothers who used illicit drugs, had mental illness, HIV infection, and others.		

Bier 2002 [2003020034]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Bayley MDI at 7-month and at 12-month	Adjusted for maternal result of Peabody Picture Vocabulary Test (PPVT) and days of oxygen	Mean Bayley MDI at 7 month, human milk group = 94 ± 7, in formula group = 90 ± 9 (Difference NS) At 12 months, human milk group = 101 ± 11, formula group = 90 ± 9 (P<0.05) ANOVA, adjusted for days of oxygen and maternal PPVT, Bayley MDI in human milk vs formula, 100 ± 12 vs 91 ± 10 (R², P<0.025) Regression analysis showed an association between the amount of milk infants received in the special care nursery (mL/wk) by gavage and/or bottle and the Bayley MDI at 7 mo (P<0.05) and 12 mo (P<0.025).	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design			x
			Confounder/bias		x	
			Blinding		x	
			Data collection			x
			Withdraw and dropout			x
			Analyses			x
			Intervention integrity			
			Actual number of mothers approached for participation was not accounted for, small sample size, cognitive test at young age, principal investigator was available to all mothers at all times to answer feeding related questions			

Eidelman 2004 [15384601]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 30.5 wk (26-33 wk) Mean BW (range): 1298 g (640-1720 g) % Male: 53% Race: Evaluated: 86 Location: Israel Sites: Single Funding:	Prospective cohort separated into 3 groups based on amount of human milk feeding; Bayley scales assessed at 6 months corrected age	<33 wk gestation; birth weight < 1750 g; exclusion criteria: intraventricular hemorrhage grade 3 or 4, perinatal asphyxia, metabolic or genetic disease	Exact amount of human milk and formula an infant received was recorded after each meal. Divided into >75% of nutrition as human milk; 25-75%; and <25%	

Eidelman 2004 [15384601]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
MDI at 6 months corrected age	Number of breastfeeding episodes was used as a covariate; MANCOVA performed with human milk and infant gender as the between subject factors	The number of breastfeeding episodes was <5%. Univariate analyses showed group differences on the MDI. Post hoc comparisons reported significant differences between the substantial milk group and the other two groups: 94.2±8.8 vs 91.7±7.2 vs 90.5±8.5 (P<0.05)	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design			x
			Confounder			x
			Blinding		x	
			Data collection			x
			Withdraw and dropout			x
			Analyses			x
			Intervention integrity			
			Not adjusted for SES, maternal education or intelligence; small sample size; cognitive testing at young age			

Elgen, 2003 [14580653]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): BW (range): <2000 g % Male: Race: Evaluate: 130 Location: Norway Sites: Single/Multi Funding:	Prospective cohort, followed 'til 11 yr	Born in the county of Hordaland, Norway, between 4/1986 and 8/1988; survived to 11 yr; low birth weight; no cerebral palsy, blindness, deafness, multiple malformations, or chromosomal abnormalities	27% received <30% human milk in neonatal ward	

Elgen, 2003 [14580653]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Weschler Preschool and Primary Scale of Intelligence- Revised at 5 yr; Weschler Intelligence Scale for Children-Revised at 11 yr	Breast milk, birth weight, paternal and maternal education, chorioamnionitis, gestational age, length of oxygen treatment, sex, Apgar at 5 minutes, smoking during pregnancy, intrapartum stress, cerebral hemorrhage in subjects with birth weight <1500 g	Unadjusted regression: lack of breast milk was associated with a mean reduction in IQ of 5.8 points (95% CI -11 to -1). After adjustment for parental confounding, breast milk was no longer a statistically significant predictor of IQ (paternal education was a significant confounder with breast milk).	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design			x
			Confounder/bias		x	
			Blinding		x	
			Data collection			x
			Withdraw and dropout			x
			Analyses		x	
			Intervention integrity			
			Lack of detailed information on breast milk exposure			

Feldman, 2003 [12918090]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 30 wk (26-33) Mean BW (range): 1298 (640-1720) % Male: Race: Enrolled/Evaluate: 86 Location: Israel Sites: Single/Multi Funding:	Prospective cohort; consecutive enrollment (6 declined)	<33 wk gestation; <1750 g Excluded infants with congenital malformation, severe hemorrhage; also excluded poverty, single parenting, teenage mothering, and maternal drug use	>75%, 25-75%, and <25% nutrition from breast milk; exact amount of breast milk and formula the infant received was recorded after each meal by nursing staff	

Feldman, 2003 [12918090]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Bayley MDI at 6 months corrected age	Multivariate ANOCOVA with breast milk and infant gender as the between-subject factors and the number of breast-feeding episodes as the covariate was computed for the cognitive indices at 6 months	An overall main effect for breast milk group was found (F=3.14, P<0.05) Univariate analysis: A (n=34) >75% group MDI 94.16 ± 8.75 B (n=21) 25-75% 91.66 ± 7.20 C (n=31) <25% 90.53 ± 8.54 (F=3.34, P<0.05, A>B, C)	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design			x
			Confounder/bias			x
			Blinding			x
			Data collection		x	
			Withdraw and dropout		x	
			Analyses			x
			Intervention integrity			
			Not adjusted for SES, maternal education or intelligence; small sample size; cognitive testing at young age			

Horwood, 2001 [11124919]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): Mean BW (range): % Male: Race: Evaluated: 298/280 analyzed Location: New Zealand Sites: Single/Multi Funding:	Cross-sectional; all 413 very low birth birthweight infants who were live born in 1986 who were admitted to one neonatal unit; retrospective recall of breast milk feeding and duration of feeding; children were assessed by WISC-R (excluded 17 children who could not reliably be assessed because of sensorineural disability; 1 child was excluded because the data on breastfeeding were missing) at 7 yr	Live born very low birthweight infants	Not breastfed (n=76) Breastfed < 4 months (n=99) 4-7 months (n=46) ≥ 8 months (n=59)	

Horwood, 2001 [11124919]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
WISC-R	Controlled for perinatal, socio-demographic (family income, single/two parent family, child ethnicity), and maternal factors (age, education, smoking); the data were analyzed using multiple linear regression techniques in which the measures of verbal and performance IQ were regressed on the four level measure of breastfeeding treated as a continuous variable and significant covariate factors.	Children who were breastfed ≥ 8 months had mean verbal IQ 10.2 (SD 0.56) higher and mean performance IQ 6.2 (SD 0.35) higher than children who did not receive breast milk. After adjustment for covariates, there remained a significant association between duration of receipt of breast milk and verbal IQ, with a 6 point advantage for infants who received breast milk for ≥ 8 months compared with no breastfeeding (P<0.001).	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design			x
			Confounder/bias		x	
			Blinding			x
			Data collection			x
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity			

O'Connor, 2003 [14508214]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
GA): <33 wk BW (range): 750-1805 g % Male: Race: Enrolled/Evaluate: 463 Location: US, UK and Chile Sites: Multi Funding:	Cohort, re-analysis of a previous RCT on supplemented formulas	Could be enrolled within 72 hours of first enteral feeding initiated by the 28 th day of life; excluded congenital abnormalities that could affect growth and development, major surgery, and others	4 mutually exclusive groups: 1. predominantly human milk (see details in paper) 2. predominantly formula 3. $\geq 50\%$ of total energy as human milk 4. $< 50\%$ of total energy as human milk	

O'Connor, 2003 [14508214]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Bayley at 12 mo; MacArthur Communicative Development Inventories at 9 mo, 14 mo	Maternal intelligence and home environment	No differences in the Bayley MDI were found among feeding groups. There was a positive association between duration of breastfeeding and the Bayley MDI at 12 months corrected age (P=0.032 in full, and P=0.073 in reduced statistical models) after controlling for home environment and maternal intelligence.	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design		x	
			Confounder/bias	x		
			Blinding			x
			Data collection			x
			Withdraw and dropout			x
			Analyses		x	
			Intervention integrity			
			Cognitive testing at a young age			

Pinelli, 2003 [2003083348]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): <1500 g Mean BW (range): % Male: Race: Enrolled/Evaluate: 148/137 Location: Canada Sites: Single Funding:	Prospective cohort (consecutive enrollment); sample from a RCT on conventional breastfeeding support versus supplementary structured breastfeeding counseling; 20 totally formula fed infants were recruited as controls	Excluded: multiple births, severe congenital abnormalities; infants of non-English speaking parents	>80% breast milk vs <80% breast milk or no breast milk	

Pinelli, 2003 [2003083348]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Bayley MDI at 12 mo	Sex, SES, birth weight, maternal age	64/128 infants received breastmilk exclusively; Adjusted MDI at 6 mo (corrected age) for >80% breastmilk group: 93 (SD 16); for <80% breastmilk group: 94 (SD 15); Adjusted MDI at 12 mo (corrected age) for >80% breastmilk group: 92 (SD 15); for <80% breastmilk group: 91 (SD 12); P>0.05 for all comparisons	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design			x
			Confounder/bias		x	
			Blinding			x
			Data collection			x
			Withdraw and dropout		x	
			Analyses			x
			Intervention integrity			
			Cognitive test at a young age			

Smith, 2003 [14630603]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Median GA (range): 27.4 wk Median BW (range): 1014 g % Male: Race: Enrolled/Evaluate: 439/420 Location: US Sites: Multi Funding:	Cohort; subjects were from Developmental Epidemiology Network cohort (n=1442), 439 enrolled in follow up studies at 6 to 8 yr	Surviving children with selected ultrasound abnormalities were included (n=119), and a random sample (1:4 ratio) of infants frequency matched for gestation were included; had undergone neuropsychological assessment and for whom parental questionnaires were complete by 9/2002	Received expressed milk without progressing to direct breastfeeding, received direct breastfeeding	No human milk

Smith, 2003 [14630603]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Kaufman Assessment Battery for Children (an estimate for overall intellectual function); Peabody Picture Vocabulary Test and Clinical Evaluation of Language Fundamentals (verbal); California Children's Verbal Learning Test (memory); Wide Range Assessment of Visual Motor Abilities (visual-spatial skill) These measures were standardized with a mean of 100 points and standard deviation of 15 points.	Maternal verbal ability, Home Observation for Measurement of the Environment Inventory, SES, duration of hospitalization	Compared with nonparticipating mothers, those who participated were older, more educated, more likely to be married, be non-smokers, and have private medical insurance. 153 subjects did not receive human milk; 142 received expressed milk without progressing to direct breastfeeding; 125 received direct breastfeeding (the majority of them first received expressed breast milk); 20% of infants received breast milk >6 mo. Breast milk feedings were associated with higher unadjusted test scores for each domain of cognitive function except memory. In the regression model that included social advantage and neonatal morbidity, the adjusted score in overall intellectual function associated with direct breastfeeding was reduced to 3.6 points (95% CI, -0.3, 7.5) from 10.7 (95%CI, 7.0, 14.4). Breastfed children had an advantage in visual-motor integration measure (adjusted IQ 5.1, 95% CI, 1.0, 9.2)	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design			x
			Confounder/bias	x		
			Blinding			x
			Data collection		x	
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity			
						27% of subjects with ultrasound abnormalities; the timing of collection of breastfeeding data was not reported

Vohr, 2006 [16818526]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 26.5 wk Mean BW (range): 785 g (BM), 794 (no) % Male: 45% Race: Enrolled/Evaluate: 1159/1035 Location: US Sites: Multi Funding:	Prospective cohort; Bayley results at 18 to 22 mos corrected age; examiner blinded to feeding history, primary caretaker stayed with the child during the Bayley examination	Enrolled prospectively in the Glutamine Trial at 15 sites of Neonatal Research Network; survivors	Breast milk – received some breast milk during hospitalization; total volume received calculated per day, values were interpolated for days of the week in which the data were not collected	

Vohr, 2006 [16818526]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments																																				
Cognitive outcome of interest: Bayley Mental Development Index (MDI)	Adjustment for marital status, maternal education, age, race/ethnicity, infant's gestational age, gender, culture positive sepsis, grade 3 to 4 intraventricular hemorrhage, oxygen at 36 wks gestational age, necrotizing enterocolitis, and weight <10 th %tile	There were 775 infants in the breast milk group and 260 in the no breast milk group. 95 subjects' Bayley could not be administered successfully; these data were not included. Mothers of infants who were seen at followup were more likely to have received prenatal care than mothers of infants who were not seen. There were no differences in infant characteristics. Information on breast milk feeding was not collected after discharge from hospital. 30.6% of infants in the breast milk group were still receiving breast milk. Mothers in the breast milk group were more likely to be white, married, have a college degree and have private health insurance. Mothers who had low household income, higher parity or were black were less likely to provide breast milk feeds. Adjusted Bayley MDI 79.9 ± 18 in BM 75.8 ± 16 in noBM P=0.0709 MDI <85 421 (58.1%) in BM 168 (70.9%) in noBM P=0.0355 Multiple regression with adjusted factors=0.53 for MDI, P=0.0002 Breast milk intake was analyzed by quintile relative to the no breast milk group; for MDI, there was a 13.1 point difference between ≤ 20th quintile and >80th quintile (P<0.0044)	<table border="1"> <tr> <td>A: strong, B: moderate, C: weak</td><td>A</td><td>B</td><td>C</td></tr> <tr> <td>Selection</td><td></td><td>x</td><td></td></tr> <tr> <td>Study design</td><td></td><td></td><td>x</td></tr> <tr> <td>Confounder/bias</td><td></td><td></td><td>x</td></tr> <tr> <td>Blinding</td><td></td><td>x</td><td></td></tr> <tr> <td>Data collection</td><td></td><td></td><td>x</td></tr> <tr> <td>Withdraw and dropout</td><td></td><td>x</td><td></td></tr> <tr> <td>Analyses</td><td></td><td></td><td>x</td></tr> <tr> <td>Intervention integrity</td><td></td><td></td><td></td></tr> </table> Residual confounding from home environment cannot be ruled out; Bayley at young age; proportion of subjects with MDI <85 could not have been adjusted; unclear if income was adjusted or not (Results did not indicate that it was, but the discussion said it was); unclear if Bayley was done blinded to feeding history (not stated in Methods, but the discussion said it was)	A: strong, B: moderate, C: weak	A	B	C	Selection		x		Study design			x	Confounder/bias			x	Blinding		x		Data collection			x	Withdraw and dropout		x		Analyses			x	Intervention integrity			
A: strong, B: moderate, C: weak	A	B	C																																				
Selection		x																																					
Study design			x																																				
Confounder/bias			x																																				
Blinding		x																																					
Data collection			x																																				
Withdraw and dropout		x																																					
Analyses			x																																				
Intervention integrity																																							

Appendix C. Evidence Tables

Part II. Preterm Infant Outcomes

Necrotizing Enterocolitis

Furman 2003

Gross 1983

Lucas 1984 1990

McGuire 2001 SRMA*

Ronnestad 2005

Schanler 2005

Tyson 1983

* Systematic Review/Meta-Analysis

Furman, 2003 [UI# 12517197]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 28 wk Mean BW (range): 1056 g % Male: 57% Race: Enrolled/Evaluate: 119 Location: Cleveland Sites: Single Funding:	Prospective cohort; The study compared the effect of varying dosages of maternal milk on neonatal outcomes	Singleton, birth weight 600-1499 g, gestational age < 33 wk, no positive drug screen, major congenital anomaly, intrauterine infection, or insurmountable maternal social factors	Intravenous dextrose during the first 24 hrs; enteral intake was begun by day 2 or 3 of life, parenteral nutrition was continued until a daily enteral intake of 120 mL/kg of body weight was reached. Infants received their mother's milk in the sequence it was expressed, except that fresh rather than frozen milk was given if available. Maternal milk was fortified, and preterm infant formula was offered when the infant reached a daily oral intake of at least 110 mL/kg. Limited availability of maternal milk was the sole reason infants were fed preterm formula in addition to maternal milk.	

Furman, 2003 [UI# 12517197]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments																																				
NEC in the 4 subgroups: No maternal milk Daily 1-24 mL/kg of maternal milk Daily 25-49 mL/kg of maternal milk Daily ≥ 50 mL/kg of maternal milk	Logistic regression; adjusted for birth weight, ethnicity, and sex	0 maternal milk 3/40 1-24 mL/kg 2/29 (OR 1.15 95% CI 0.8-12.13) 25-49 mL/kg 2/18 (OR 1.99 95% CI 0.14-21.03) ≥ 50 mL/kg 0/32 (OR 0 95% CI 0 – 3.56)	<table> <tr> <td>A: strong, B: moderate, C: weak</td><td>A</td><td>B</td><td>C</td></tr> <tr> <td>Selection</td><td></td><td></td><td></td></tr> <tr> <td>Study design</td><td></td><td></td><td>x</td></tr> <tr> <td>Confounder</td><td></td><td>x</td><td></td></tr> <tr> <td>Blinding</td><td></td><td></td><td>x</td></tr> <tr> <td>Data collection</td><td></td><td>x</td><td></td></tr> <tr> <td>Withdraw and dropout</td><td></td><td>x</td><td></td></tr> <tr> <td>Analyses</td><td></td><td>x</td><td></td></tr> <tr> <td>Intervention integrity</td><td></td><td></td><td>x</td></tr> </table>	A: strong, B: moderate, C: weak	A	B	C	Selection				Study design			x	Confounder		x		Blinding			x	Data collection		x		Withdraw and dropout		x		Analyses		x		Intervention integrity			x
A: strong, B: moderate, C: weak	A	B	C																																				
Selection																																							
Study design			x																																				
Confounder		x																																					
Blinding			x																																				
Data collection		x																																					
Withdraw and dropout		x																																					
Analyses		x																																					
Intervention integrity			x																																				

Gross, 1983 [UI# :6848932]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 31 wk Mean BW (range): 1322 g % Male: Race: Enrolled/Evaluate: Location: US Sites: Single Funding:	Random number assignment into 3 groups: mature human milk, preterm human milk, whey-based formula; any infant withdrawn from the study was replaced by the next infant enrolled, until there were 20 infants in each group; remained in the study until they reached 1800 g	27 – 33 wk gestation; $\leq$ 1600 g; no congenital abnormalities or major diseases; ability to begin enteral feed by 6 th day of life; absence of breast milk from own mother	Feedings began between the first and sixth days of life.	

Gross, 1983 [UI#:6848932]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
NEC (abdominal distention, gastric aspirates, guaiac-positive stools, and pneumatosis intestinalis)	NEC was analyzed as infants who withdrew from the study	67 infants recruited for the study; 1/41 (BM) vs. 3/26 (FM) with NEC	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design	x		
			Confounder			x
			Blinding		x	
			Data collection		x	
			Withdraw and dropout		x	
			Analyses			x
			Intervention integrity			

Lucas, 1984 [UI# 6476868]; 1990 [UI# 1979363]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 31 wk Mean BW (SD): 1370 g (320 g) % Male: Race: Enrolled/Evaluate: Location: UK Sites: /Multi Funding:	Randomization within 48 h of birth for infants whose mothers decided not to breastfeed, randomized into banked donor human milk vs. preterm formula; only infants fed exclusively on donor milk or formula were compared	<1850 g, excluded severe congenital abnormalities		

Lucas, 1984, [UI# 6476868]; 1990 [UI# 1979363]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Confirmed NEC – required surgery or who died	Logistic regression was used to adjust for factors that have been associated with NEC.	1/86 (BM) vs. 4/76 (FM) OR 4.7 (0.5 – 43)	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design	x		
			Confounder	x		
			Blinding		x	
			Data collection	x		
			Withdraw and dropout	x		
			Analyses	x		
			Intervention integrity			

Evidence table for Systematic Reviews of Breastfeeding and NEC

Author, Year[UI]	Topic	McGuire, 2001 [11687169]	NEC
Literature search (Dates)1966-10/2003		Medline; Other databases searched? (yes, CENTRAL, EMBASE, CINAHL); unpublished data used? (no)	
Countries where primary studies conducted		ND	
Study eligibility / inclusion criteria		RCT comparing feeding with formula milk versus term human milk in low birth weight or preterm infants	
Study design [No. Of studies]		RCT: 3 studies with outcomes re: NEC	
No. of subjects		287	
Study population (definition in included studies)		Gross 1983 27-33 wk, bw < 1600 g Tyson 1983 very low birth weight infants Lucas 1984 preterm infants < 1850 g	
Intervention/Exposure (definition in included studies)		Gross 1983 unfortified term donor breast milk, fed until 1800 g or until withdrawal secondary to feed intolerance or NEC Tyson 1983 pooled banked term human milk, allocation at 10 th day of life, fed until 2000 g or until withdrawal secondary to illness requiring parenteral fat or protein Lucas 1984 banked term human milk 200 mL/kg/d, fed until 2000 g or until d/c	
Comparator (definition in included studies)		Gross 1983 standard calorie (0.67 kcal/mL, protein enriched (1.9 g/100 mL) formula Tyson 1983 enriched calorie (0.87 kcal/mL), protein enriched (2.2 g/100 mL) formula Lucas 1984 enriched calorie (0.80 kcal/mL), protein enriched (2.0 g/100 mL) formula at 180 mL/kg/d	
Outcomes (definition in included studies)		Necrotizing enterocolitis	
Heterogeneity assessments			
Quality assessments		Gross 1983 & Tyson 1983 The studies did not attempt blinding of parents, carers, or assessors. Adverse events were not reported as primary end points but rather as withdrawal criteria. Lucas 1984 The study did not attempt blinding of parents, carers, or assessors. Data on NEC on all 159 recruited infants were reported	
Publication bias assessments			
Statistical Analysis or meta-analytic methods		3 trials were combined in a meta-analysis: formula vs human milk, RR: 2.5 (95% CI 0.9, 7.3); RD: 0.05 (95% CI 0.00, 0.1)	
Results		Gross 1983 incidence of NEC in formula vs human milk: RR 2.4 (95% CI 0.3, 21.6); RD: 0.07 (95% CI 0.09, 0.22) Tyson 1983 incidence of NEC in formula vs human milk: RR 4.2 (95% CI 0.2, 85.3); RD: 0.05 (95% CI -0.03, 0.12) Lucas 1984 incidence of NEC in formula vs human milk: RR 2.2 (95% CI 0.6, 8.4); RD: 0.04 (95% CI -0.03, 0.12)	
Quality of the systematic review			
Author's interpretations of the results		We did not find any statistically significant difference in the risk of NEC with formula milk versus term human milk feeding.	
Comments / Limitations		Donor milk, not mother specific milk	

Author, yr: McGuire, 2001

Topic of the systematic review*/MA:

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Y
2	Types of exposure or intervention used	Y
3	Types of study designs used	Y
4	Study population	Y
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Y
6	Effort to include all available studies (contact with authors)	Y
7	Databases and registries searched	Y
8	Method of addressing published articles other than English	Y/N
9	Method of handling abstracts and unpublished studies	N
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Y
11	Assessment of confounding	Y/N
12	Assessment of quality	Y
13	Assessment of heterogeneity	Y
14	Description of statistical methods sufficient to replicate	Y
15	Provision of appropriate tables and graphs	Y
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Y
17	Tables giving descriptive information of each study included	N
18	Results of sensitivity testing (subgroup analysis)	Y
19	Indication of statistical uncertainty of findings	Y
20	Quantitative assessment of bias (eg. publication bias)	N
	Overall quality	

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Y
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	Y
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Y
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	Y/N
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Y/N
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Y/N
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Y
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Y
8	Were results reported accurately (eg, no discrepancies between text and tables)?	Y/N
9	Were conclusions justified by the reported/collected data and analysis?	Y

Ronnestad, 2005 [UI# 15687416]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 22-27 wk Mean BW (range): 845 g (est.) % Male: 56% Race: Enrolled/Evaluate: 462 enrolled; 405 survived until day 7 Location: Norway Sites: Multi-center Funding:	Prospective cohort	Participating centers had a common policy of achieving full enteral feeding with the mother's milk or banked donor milk as early as possible, although there was no consensus in terms of a detailed protocol for feeding strategies. All infants born in Norway in 1999 and 2000 with gestational age of <28 wk or birth weight < 1000 g.	Tube feeding with human milk was usually started within a few hours after delivery, with 1 to 2 mL of milk every 2 or 3 hrs, increasing by 0.5 to 1 mL every 6 to 8 hours as tolerated. Enteral nutrition was supplemented with parenteral glucose from day 1, amino acids and lipids from day 2 and day 3, respectively.	

Ronnestad, 2005 [UI# 15687416]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
This was a study primarily examining late-onset septicemia (after day 6 of life) that also reported outcome on NEC after day 6 of life.	No analysis/adjustment specific to NEC was reported.	Enteral feeding with human milk was commenced within 1, 2, or 3 days for 61%, 92%, and 96% of the infants. 9/405 (2.2%) patients had confirmed NEC.	A: strong, B: moderate, C: weak	A	B	C
			Selection			
			Study design			x
			Confounder			x
			Blinding			x
			Data collection			
			Withdraw and dropout		x	
			Analyses			x
			Intervention integrity			x

Schanler, 2005 [UI# 16061595]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 23-26wk & 27-29 wk Mean BW (range): 952 g % Male: 53% Race: ND Enrolled/Evaluate: 243 Location: Texas Children's Hospital Sites: Multi Funding: government	RCT 4 days to 90 days of age or discharge from the hospital	Infants with mother expected to breastfeed; if mother's own milk was unavailable; then infants were randomized to Donor Milk versus Preterm formula	Donor milk (DM) (+ mother's milk partially) Reference group: non-randomized, exclusive mother's milk (MM) Small quantities of mother's milk (~20 mL/kg/day) was initiated in the first week and continued for ~3 to 5 days before the volume was advanced. Milk intake was increased by ~20 mL/kg daily to 100 mL/kg, at which time milk fortifier was added, this was advanced by ~20 mL/kg daily until 160 mL/kg per day was achieved.	Preterm formula (PF) (+ mother's milk partially)

Schanler, 2005 [UI# 16061595]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments																																				
NEC Only cases of NEC that occurred after the infant attained a milk intake of ≥ 50 mL/kg were analyzed for the primary outcome.	17 infants (21%) in DM were switched to PF because of poor weight gain. All infants remained in the original assigned group for analysis.	Incidence of NEC: DM vs PF: 5/78 vs 10/88 ($P=0.27$) Non-randomized group MM had fewer repeated episodes of late-onset-sepsis and/or NEC (OR 0.18; 95% CI 0.04–0.79) compared with combined groups DM and PF	<table> <tr> <td>A: strong, B: moderate, C: weak</td><td>A</td><td>B</td><td>C</td></tr> <tr> <td>Selection</td><td></td><td>x</td><td></td></tr> <tr> <td>Study design</td><td>x</td><td></td><td></td></tr> <tr> <td>Confounder</td><td>x</td><td></td><td></td></tr> <tr> <td>Blinding</td><td></td><td>x</td><td></td></tr> <tr> <td>Data collection</td><td></td><td></td><td></td></tr> <tr> <td>Withdraw and dropout</td><td></td><td>x</td><td></td></tr> <tr> <td>Analyses</td><td></td><td>x</td><td></td></tr> <tr> <td>Intervention integrity</td><td></td><td>x</td><td></td></tr> </table>	A: strong, B: moderate, C: weak	A	B	C	Selection		x		Study design	x			Confounder	x			Blinding		x		Data collection				Withdraw and dropout		x		Analyses		x		Intervention integrity		x	
A: strong, B: moderate, C: weak	A	B	C																																				
Selection		x																																					
Study design	x																																						
Confounder	x																																						
Blinding		x																																					
Data collection																																							
Withdraw and dropout		x																																					
Analyses		x																																					
Intervention integrity		x																																					

Tyson, 1983 [UI# - 6864403]

Study characteristics	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean GA (range): 31.5 wk Mean BW (range): 1232 g % Male: Race: Enrolled/Evaluate: Location: US Sites: Single Funding:	RCT; infants were not enrolled until the 10 th day; enrolled into donor human milk or formula milk groups	Infants whose mothers continued to provide milk at 10 days were excluded, as infants with any significant illnesses		

Tyson, 1983 [UI# - 6864403]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
NEC confirmed at surgery; suspected NEC	NEC was analyzed as subjects who dropped out of the study.	0/37 (BM) vs. 1/44 (FM) for confirmed NEC	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design		x	
			Confounder			x
			Blinding		x	
			Data collection		x	
			Withdraw and dropout		x	
			Analyses			x
			Intervention integrity			

Appendix C. Evidence Tables

Part III. Maternal Outcomes

Breast Cancer

Bernier 2000 SRMA*
CGHFBC 2002 SRMA*
Gammon 2002
Jernstrom 2004
Lee 2003
Lipworth 2000 SRMA*

* Systematic Review/Meta-Analysis

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI]	Topic	Bernier, 2000	Breast cancer
Literature search (Dates)		Medline (1980-1998); Other databases searched? (yes-Embase); unpublished data used? (no)	
Countries where primary studies conducted		Developed and developing countries (Estonia, Canada, USA, Sicilia, Costa Rica, Australia, India, China, Greece, Italy, Mexico, New Zealand)	
Study eligibility / inclusion criteria		A study was included in the analysis: 1) it was published between 1980 and 1998 2) it presented primary original data; 3) it was published either in English or French; 4) the type of study was either a case-control, or a prospective cohort study; 5) data on breastfeeding exposure were reported and used in the analysis 6) breast cancer was the event of interest 7) specific odds ratio measuring the breastfeeding and breast cancer association were given or could be derived. Excluded were studies in which data from parous women who had not breastfed could not be separated from nulliparous women	
Study design [No. Of studies]		23 [Case-control studies]	
No. of subjects		25871 cases/44910 controls	
Study population (definition in included studies)		Parous women with breast cancer	
Intervention/Exposure (definition in included studies)		Breastfeeding (Ever)	
Comparator (definition in included studies)		Breastfeeding (never)	
Outcomes (definition in included studies)		Breast cancer (histologically diagnosed)	
Heterogeneity assessments		Yes; homogeneity graph and by subgroup analysis (menopausal status; duration of breastfeeding)	
Quality assessments		Unclear (MA discussed methodologic quality with subheading of origin of the women; breast cancer diagnosis; and data collection)	
Publication bias assessments		Yes (funnel plot)	
Statistical Analysis or meta-analytic methods		Both random and fixed effects model	
Results		Breast cancer risk ever vs never breastfeeding Pooled OR using the fixed effects model 0.90 (0.86-0.94) Pooled OR using random effects model 0.84 (0.78-0.91) These results were not modified using adjusted ORs Breast cancer risk according to the menopausal status at the time of the diagnosis in ever vs never breastfeeding mothers Menopausal at the time of breast cancer diagnosis Pooled OR using the fixed effects model 1.01 (0.94-1.09) Pooled OR using random effects model 0.84 (0.69-1.03) Non-menopausal at the time of breast cancer diagnosis Pooled OR using the fixed effects model 0.83 (0.76-0.91) Pooled OR using random effects model 0.81 (0.72-0.91) Breast cancer risk according to duration of breast feeding 0+ to 6 mo v never Pooled OR using the fixed effects model 1.02 (0.94-1.10) Pooled OR using random effects model 1.00 (0.86-1.16) 6+ to 12 mo v never Pooled OR using the fixed effects model 0.97 (0.88-1.06) Pooled OR using random effects model 0.97 (0.86-1.09) 12+ mo vs. never Pooled OR using the fixed effects model 0.75 (0.70-0.80)	

Author, Year[UI] Topic	Bernier, 2000 Breast cancer
	Pooled OR using random effects model 0.72 (0.65-0.80)
Quality of the systematic review	B
Author's interpretations of the results	Breastfeeding appeared to be a protective factor but was of small magnitude compared with other known risk factors of breast cancer
Comments / Limitations	Quality score for individual studies not assessed

Author, yr: Bernier 2000

Topic of the systematic review*/MA: Breast cancer and breastfeeding

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Y
2	Types of exposure or intervention used	Y
3	Types of study designs used	Y
4	Study population	Y
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Y
6	Effort to include all available studies (contact with authors)	N
7	Databases and registries searched	Y
8	Method of addressing published articles other than English	N
9	Method of handling abstracts and unpublished studies	N
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Y
11	Assessment of confounding	Y
12	Assessment of quality	Unclear
13	Assessment of heterogeneity	Y
14	Description of statistical methods sufficient to replicate	Y
15	Provision of appropriate tables and graphs	Y
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Y
17	Tables giving descriptive information of each study included	Partially yes
18	Results of sensitivity testing (subgroup analysis)	Y
19	Indication of statistical uncertainty of findings	N
20	Quantitative assessment of bias (eg. publication bias)	Y
	Authors' Conclusions	
21	Consistent with the presented results	Y
22	Commenting on the validity of the conclusions	Y
23	Presenting alternative explanations	Y
	Overall quality	B

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	Yes
2	Did authors give justifications for inclusion/exclusion criteria pertaining to studies of interest?	Partially
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	Yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	Partially
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	No
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	Yes
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	Yes
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	Yes
8	Were results reported accurately (eg, no discrepancies between text and tables)?	No
9	Were conclusions justified by the reported/collected data and analysis?	Yes

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI] Topic	Collaborative Group on Hormonal Factors in Breast Cancer (CGHFBC) 2002 [12133652] Breast Cancer
Literature search (Dates)	ND Available in prior publications Medline (nd); Other databases searched? (nd); unpublished data used? (yes)
Countries where primary studies conducted	International studies including developed and developing nations
Study eligibility / inclusion criteria	Studies that had data on at least 100 women with incident invasive breast cancer and had recorded information on each women with respect to reproductive factors and use of hormonal preparations. For the cohort studies, a nested case-control study design was used, in which four randomly selected controls per case were matched for age at diagnosis and where appropriate, broad geographical region
Study design [No. Of studies]	case-control and cohort studies [47]
No. of subjects	50,302/96,973
Study population (definition in included studies)	Women with breast cancer and controls
Intervention/Exposure (definition in included studies)	Breastfeeding (Ever and lifetime months of breastfeeding (median))
Comparator (definition in included studies)	Breastfeeding Never
Outcomes (definition in included studies)	Breast cancer (ND)
Heterogeneity assessments	Subgroup analyses including women from developed and developing countries, women of different ages, ethnic origins, familial patterns of disease, and 11 other possible relevant factors
Quality assessments	None
Publication bias assessments	ND
Statistical Analysis or meta-analytic methods	Analyses were routinely stratified by study, by center within study, by fine divisions of age, by age at first birth and by menopausal status. Where appropriate, parous women are further stratified by fine divisions of parity. Data combined by means of the Mantel-Haenszel stratification technique, the stratum specific quantities calculated being the standard observed minus expected numbers of women with breast cancer, together with their variances and covariates.
Results	% reduction in relative risk of breast cancer per 12 months of breastfeeding and 99% (SE)= 4.3 (0.8) Data also available for floating absolute risk and floated SE for subgroups Parous women who never breastfed v ever breastfed Life time months of breastfeeding
Quality of the systematic review	A
Author's interpretations of the results	The longer women breastfeed the more they are protected against breast cancer. The lack of or short lifetime duration of breastfeeding typical of women in developed countries makes a major contribution to the high incidence of breast cancer in these countries.
Comments / Limitations	Individual patient level data

Author, yr: Collaborative group on Hormonal factors in breast cancer, 2002

Topic of the systematic review*/MA: Breast cancer and breastfeeding

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Y
2	Types of exposure or intervention used	Y
3	Types of study designs used	Y
4	Study population	Y
	Reporting of Search strategy	
5	Search strategy, including time period and key words	N
6	Effort to include all available studies (contact with authors)	N
7	Databases and registries searched	N
8	Method of addressing published articles other than English	N
9	Method of handling abstracts and unpublished studies	N
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Y
11	Assessment of confounding	Y
12	Assessment of quality	N
13	Assessment of heterogeneity	Y
14	Description of statistical methods sufficient to replicate	Y
15	Provision of appropriate tables and graphs	Y
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Y
17	Tables giving descriptive information of each study included	Y
18	Results of sensitivity testing (subgroup analysis)	Y
19	Indication of statistical uncertainty of findings	N
20	Quantitative assessment of bias (eg. publication bias)	N
	Authors' Conclusions	
21	Consistent with the presented results	Y
22	Commenting on the validity of the conclusions	Y
23	Presenting alternative explanations	Y
	Overall quality	A

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	No
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	yes
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	partially
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Individual patient level data
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	yes
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	yes

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	yes
8	Were results reported accurately (eg, no discrepancies between text and tables)?	yes
9	Were conclusions justified by the reported/collected data and analysis?	yes

Gammon 2002 [12206514]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): <35 to 85+ yrs Menopause: mix Race: white 92-93%; Black 5%; and other 2-3% Enrolled/Evaluate: 1508 cases/1556 controls Location: population based Sites: Multi Funding: Gov	Women with newly diagnosed breast cancer (invasive and in-situ)	Case control study design Follow-up duration nd	Cases were women newly diagnosed with a first primary <i>in situ</i> or invasive breast cancer between August 1, 1996, and July 31, 1997, confirmed by the physician and the medical record, who were residents of either Nassau or Suffolk counties in New York at the time of their diagnosis, and who spoke English. (Over 97% of all residents in these two counties are English-speaking.) Controls: Control women were a sample of current residents of Nassau and Suffolk counties who spoke English, who did not have a personal history of breast cancer, and who were frequency matched to the expected distribution of case women by 5-year age group.	By duration: <2 months 2-5 months 6-13 months 14+ months	Never lactated

Gammon 2002 [12206514]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Risk for breast cancer among those who donated a blood sample	Unconditional and multiple logistic regression model identified using backward selection model	Among parous women lactation for 14 or more months decreased the odds of breast cancer (n=102) 0.70 (0.54, 0.92) compared to those who never lactated (ref group) or less than lactated <14 mo Lactation <2 mo, 2-5 mo, and 6-13 mo when compared to never lactated group had non significant results for the odds of breast cancer development.	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design		x	
			Confounder			x
			Blinding			x
			Data collection	x		
			Withdraw and dropout			
			Analyses			x
			Intervention integrity			x
			Overall: C			

Jernstrom 2004 [15265971]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 39 Menopause: ND Race: Jewish 30% cases French Canadian 10% Other white 59% Black 1% Other 0.5% Enrolled/Evaluate: 1927/1830 Location: hospital based Sites: Multi Funding: ND	Women with carriers of gene mutation BRCA1 and BRCA2	Case-control study design Follow-up 8.2 (time between breast cancer diagnosis and the completion of study questionnaire.	Selected Eligible subjects who carried a mutation in either the BRCA1 or the BRCA2 gene were drawn from a database	For women who gave an approximate length of breastfeeding (e.g., 3–5 months), the midpoint of duration was used. The year of breast cancer diagnosis and years of all pregnancies were recorded. If both occurred during the same calendar year, it was not possible to establish the sequence of breast cancer and pregnancy. Only live births that occurred at least one calendar year before a diagnosis of breast cancer were captured. Twin births were considered to be one birth with respect to breast-feeding.	Exposure information on control subjects (e.g., breast-feeding, parity) was restricted to the period before the date of a diagnosis of breast cancer in the matched case subject.

Jernstrom 2004 [15265971]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Breast cancer among BRCA1 and BRCA 2	Conditional logistic regression for odds ratios (ORs) in the univariate and multivariate analyses. In the first analysis, parity and oral contraceptive use were used as covariate variables. Because of the strong association between parity and breast-feeding, secondary analysis was performed a that excluded all nulliparous women. All analyses were adjusted for oral contraceptive use.	In univariate analyses subjects who had breastfed their kids >1y and had BRCA1 gene polymorphism (OR 0.56; 95% CI 0.41 to 0.76) had statistically significant reduced odds of breast cancer compared to none breastfeeding. Breastfeeding ≤1 yr + BRCA1 (OR 0.89; 95% CI 0.70, 1.15; p<.001) and breastfeeding + BRCA2 were not associated with reduced odds. In multivariate subjects who had breastfed their kids >1y and had BRCA1 gene polymorphism (OR 0.55; 95% CI 0.38 to 0.80) had statistically significant reduced odds of breast cancer compared to none breastfeeding. The trend per month of breast feeding also decreased the odds of breast cancer	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design		x	
			Confounder	x		
			Blinding			
			Data collection		x	
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity		x	
			Overall grade: B			

Lee SY, 2003 [12704674]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 33-44 Menopause: Pre Race: Asian 100% Enrolled/Evaluate: 110604 Location: population based Sites: Multi Funding: Gov	Earlier menarche Oral contraceptive use Ever smoke Alcohol use Exercise (yes) [data available by duration of lactation]	Prospective observational cohort Follow-up 6 yrs	Premenopausal women who were enrolled in the Korean Medical Insurance Corporation (KMIC) and participated in the Korean Women's cohort	Ever vs Never Periods of lactation 1-12 months 13-24 months	Ever vs Never Periods of lactation 1-12 months 13-24 months

Lee SY, 2003 [12704674]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Incident breast cancers	In bivariate analyses, the authors examined the relation between lactation and traditional breast cancer risk factors, adjusting for age. In these bivariate analyses, the authors tested for trends across categories of duration of lactation, using parous women who never breastfed as the reference. In Cox proportional hazard models, age, age at menarche, number of children, age at first pregnancy, oral contraceptives, smoking, exercise and obesity were controlled to delineate the independent effects of lifetime lactation. In all analyses, a 2-sided alpha level of 0.05 was considered statistically significant.	Compared to parous women who had no history of lactation, periods of lactation of 1–12 months or 13–24 months were associated with decreased risk of breast cancer (RR, 0.8; 95% CI, 0.7–1.0 or RR, 0.7; 95% CI, 0.5–1.1, respectively), and this risk decreased even further for those with who breastfed for more than 24 months (RR, 0.6; 95% CI, 0.3–1.0)				A: strong, B: moderate C: weak	A	B	C
						Selection		x	
						Study design	x		
						Confounder	x		
						Blinding			x
						Data collection			x
						Withdraw and dropout		x	
						Analyses		x	
						Intervention integrity		x	
							Overall: B		

Evidence table for Systematic Reviews of Breastfeeding

Author, Year[UI]	Topic	Lipworth, 2000 [10675379]	Breast Cancer
Literature search (Dates)		Medline (1966-1998); Other databases searched? (no); unpublished data used? (no)	
Countries where primary studies conducted		International studies including developed and developing nations	
Study eligibility / inclusion criteria		Studies that included more than 200 cases overall and explicitly controlled for number of full-term pregnancies and age at first birth	
Study design [No. Of studies]		Case-control studies	
No. of subjects		23461/52948	
Study population (definition in included studies)		Premenopausal women with breast cancer and controls (nd)	
Intervention/Exposure (definition in included studies)		Breastfeeding (ever)	
Comparator (definition in included studies)		Breastfeeding (never)	
Outcomes (definition in included studies)		Breast cancer (nd)	
Heterogeneity assessments		None	
Quality assessments		ND	
Publication bias assessments		ND	
Statistical Analysis or meta-analytic methods		No meta-analysis; only qualitative review	
Results		"Ever" breastfeeding effect remains inconclusive with results indicating either no association or a weak protective effect against breast cancer. Few studies also reported an inverse association between increasing life-time breastfeeding and breast cancer risk in parous women.	
Quality of the systematic review		B	
Author's interpretations of the results		"Evidence supporting protective factor of breastfeeding for breast cancer is limited and should be interpreted with caution."	
Comments / Limitations			

Author, yr: Lipworth, 2000 [10675379]

Topic of the systematic review*/MA: Breast Cancer

		Reporting yes/no
	Reporting of Background	
1	Description of study outcomes	Yes
2	Types of exposure or intervention used	Yes
3	Types of study designs used	Yes
4	Study population	Yes
	Reporting of Search strategy	
5	Search strategy, including time period and key words	Yes
6	Effort to include all available studies (contact with authors)	Partially
7	Databases and registries searched	Partially
8	Method of addressing published articles other than English	No
9	Method of handling abstracts and unpublished studies	No
	Reporting of Methods	
10	Rationale for selection and/or coding of data	Yes
11	Assessment of confounding	Yes
12	Assessment of quality	No
13	Assessment of heterogeneity	Yes
14	Description of statistical methods sufficient to replicate	Not applicable
15	Provision of appropriate tables and graphs	Partially
	Reporting of Results	
16	Graphic summarizing of individual study estimates and overall	Not applicable
17	Tables giving descriptive information of each study included	Yes
18	Results of sensitivity testing (subgroup analysis)	Not applicable
19	Indication of statistical uncertainty of findings	Not applicable
20	Quantitative assessment of bias (eg. publication bias)	Not applicable
	Overall quality	B

*When grading a systematic review, please skip item #14, 18-20

Please fill the white areas only

	Evaluation of quality of a systematic review or meta-analysis	Yes /No /Partially
1	Is search strategy appropriate/relevant to the key question(s)?	yes
2	Did authors give justifications for inclusion/exclusion criteria pertaining to the studies of interest?	no
3	Were Population, Intervention/Exposure, Comparator/Control, Outcomes, and Study Designs well-defined?	yes
4	Did authors attempt to minimize errors in data extraction (such as collecting the relevant data, double data extraction, verification of data extraction...etc.)?	nd
5	Was individual study quality (such as sample size, study design, blinding, various biases and confounders, study subject attrition rate...etc.) assessed?	Partially
6	Did authors consider appropriate confounders and justification for adjusting or not adjusting for those confounders?	yes
7	Was the justification for combining or not-combining data for meta-analysis, or the use of a particular analytic method or model appropriate?	no
8	Was heterogeneity explored (this is in addition to any statistical test of heterogeneity, did the authors look at the actual differences in PICOS, for example?)?	no
8	Were results reported accurately (eg, no discrepancies between text and tables)?	no
9	Were conclusions justified by the reported/collected data and analysis?	yes

Appendix C. Evidence Tables

Part III. Maternal Outcomes

Osteoporosis-BMD

Hansen 1991
Matsushita 2002
Sowers 1992
Uusi-Rasi 2002

Osteoporosis-Fracture

Alderman 1986
Chan 1996
Cumming 1993
Hoffman 1993
Kreiger 1982
Michaelsson 2001

Hansen, 1991 [UI#1790399]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 51 Menopause: post Race: ND Enrolled/Evaluate: 178/121 Location: Denmark Sites: Single Funding:	healthy	Prospective cohort (12 years)	In 1979, 178 women completed a 2-year, prospective study receiving placebo. They had been selected by questionnaire and a medical screening examination. All had passed a natural menopause 6 months to 3 years earlier and were free of disease or medications known to influence the calcium metabolism. In 1989, 12 years after the start of the original study, all 178 women were re-contacted for a follow-up examination. Of these, 24 (13%) were unable to participate: 6 women had died, 9 suffered from diseases which made participation impossible, 6 unwilling to participate for personal reasons, 2 had emigrated, and 1 could not be located. Among the 154 women underwent the follow-up examination, 24 women were excluded due to owing to post-trial treatment with sex hormones for more than 1 year and another 9 women presenting with disease known to influence calcium metabolism were excluded as well.	111 women (92%) had experienced 1 or more pregnancies (mean: 2.6); 87% breast fed for a mean total lactation period of 13 months	No lactation

Hansen, 1991 [UI#1790399]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments																																																
Lumbar spine BMD, measured by DEXA. BMD is calculated as BMC divided by the area of interest from L2 to L4 including the intervertebral discs. BMD only measured in the follow-up. Forearm BMC was measured by single photon absorptiometry (¹²⁵I). The rate of postmenopausal bone loss (% per year) was determined as the difference in % between the measurement in 1979 and the measurement obtained in 1989 divided by 10 years.	Student's t test for unpaired data	<table><tr><td></td><td>$\Delta\text{BMC}_{\text{arm}}\%$ per year</td><td>$\Delta\text{BMD}_{\text{spine}}$</td></tr><tr><td>Lactation (-)</td><td>-1.8±0.8</td><td>0.84±0.18</td></tr><tr><td>Lactation (+)</td><td>-1.7±0.8</td><td>0.88±0.14</td></tr><tr><td>p-value</td><td>NS</td><td>NS</td></tr></table>		$\Delta\text{BMC}_{\text{arm}}\%$ per year	$\Delta\text{BMD}_{\text{spine}}$	Lactation (-)	-1.8±0.8	0.84±0.18	Lactation (+)	-1.7±0.8	0.88±0.14	p-value	NS	NS	<table><tr><td>A: strong, B: moderate, C: weak</td><td>A</td><td>B</td><td>C</td></tr><tr><td>Selection</td><td></td><td></td><td>x</td></tr><tr><td>Study design</td><td></td><td>x</td><td></td></tr><tr><td>Confounder</td><td></td><td></td><td>x</td></tr><tr><td>Blinding</td><td></td><td>x</td><td></td></tr><tr><td>Data collection</td><td></td><td>x</td><td></td></tr><tr><td>Withdraw and dropout</td><td></td><td></td><td>x</td></tr><tr><td>Analyses</td><td></td><td></td><td>x</td></tr><tr><td>Intervention integrity</td><td></td><td></td><td></td></tr></table> Overall: C Univariate analyses for BMD/BMC outcomes and lactation. Unclear whether no lactation group including non-parous women	A: strong, B: moderate, C: weak	A	B	C	Selection			x	Study design		x		Confounder			x	Blinding		x		Data collection		x		Withdraw and dropout			x	Analyses			x	Intervention integrity			
	$\Delta\text{BMC}_{\text{arm}}\%$ per year	$\Delta\text{BMD}_{\text{spine}}$																																																	
Lactation (-)	-1.8±0.8	0.84±0.18																																																	
Lactation (+)	-1.7±0.8	0.88±0.14																																																	
p-value	NS	NS																																																	
A: strong, B: moderate, C: weak	A	B	C																																																
Selection			x																																																
Study design		x																																																	
Confounder			x																																																
Blinding		x																																																	
Data collection		x																																																	
Withdraw and dropout			x																																																
Analyses			x																																																
Intervention integrity																																																			

Matsushita, 2001 [UI#12200655]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 30.1 Menopause: pre Race: Asian (assumed) Enrolled/Evaluate: 113/113 Location: Japan Sites: Single Funding: ND	Healthy, well-nourished	Longitudinal cohort study Mean interval between each delivery = 2.5 years	Postpartum Japanese women, aged 20-40 years, who delivered a singleton infant at the university hospital between 1994 and 1999. Women who had 2 nd pregnancy were included. Exclusion criteria included complications and medications before or during pregnancy, that would affect bone metabolism, delivery before 36 weeks' gestation, or bed rest in excess of 7 days.	Lactation information including length and source of lactation (breast-fed or formula-fed) was collected by questionnaire immediately after the subsequent delivery. Only 1 woman had formula-fed, and the other breast fed with (n=46) or without (n=66) formula fed. Whether breast-feeding had been exclusive, predominant, or on-demand, and when alternative food had been introduced, was not available because of ambiguous memory.	

Matsushita, 2001 [UI#12200655]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments																																								
BMD of the lumbar spine (L2-L4) was measured by DEXA within 7 days of delivery. 113 of women became pregnant again and they had 2 nd BMD measurement within 7 days of the next delivery. The percent change in BMD (Δ BMD%) was calculated by subtracting the value at the time of the first pregnancy from the value at the time of the second pregnancy.	BMD and backgrounds of consecutive pregnancies were compared using a paired t-test. Multiple regression analysis was performed to look for independent predictors of Δ BMD%. BMD after the initial delivery was included as an independent variable to adjust for regression toward the mean.	Lumbar BMD of 110 women after the initial and the subsequent delivery was 1.006 ± 0.117 and 1.019 ± 0.115 g/cm ² , respectively. BMD after the subsequent delivery was significantly higher than that after the initial delivery ($p=0.001$), with a Δ BMD% of $1.4 \pm 4.2\%$. Independent determinants of Δ BMD% were explored by multiple regression analysis: the length of lactation between the scans showed no correlation with Δ BMD% (coefficient= -0.06 ± 0.157 SEM, $p=0.702$). Age was the most significant predictor for Δ BMD% in the model.	<table border="1"> <tr> <td>A: strong, B: moderate, C: weak</td> <td>A</td> <td>B</td> <td>C</td> </tr> <tr> <td>Selection</td> <td>x</td> <td></td> <td></td> </tr> <tr> <td>Study design</td> <td>x</td> <td></td> <td></td> </tr> <tr> <td>Confounder</td> <td>x</td> <td></td> <td></td> </tr> <tr> <td>Blinding</td> <td></td> <td>x</td> <td></td> </tr> <tr> <td>Data collection</td> <td>x</td> <td></td> <td></td> </tr> <tr> <td>Withdraw and dropout</td> <td>x</td> <td></td> <td></td> </tr> <tr> <td>Analyses</td> <td></td> <td>x</td> <td></td> </tr> <tr> <td>Intervention integrity</td> <td></td> <td></td> <td></td> </tr> <tr> <td>Overall: A</td> <td></td> <td></td> <td></td> </tr> </table>	A: strong, B: moderate, C: weak	A	B	C	Selection	x			Study design	x			Confounder	x			Blinding		x		Data collection	x			Withdraw and dropout	x			Analyses		x		Intervention integrity				Overall: A			
A: strong, B: moderate, C: weak	A	B	C																																								
Selection	x																																										
Study design	x																																										
Confounder	x																																										
Blinding		x																																									
Data collection	x																																										
Withdraw and dropout	x																																										
Analyses		x																																									
Intervention integrity																																											
Overall: A																																											

Sowers, 1992 [UI#4543]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Menopause: mix Race: European Enrolled/Evaluate: 217/181 Location: US Sites: Single Funding: Government	ND	Prospective cohort (5 year)	In 1984, BMD of women In aged 22-54 years living in a single rural community were measured by single-photon densitometry. The community was selected for study because of the mineral characteristics of the community water supply. Women were eligible to participate in the baseline study if they had lived in the community the previous 5 years, were able to complete face-to-face interview, were ambulatory and reported themselves to be of northern European heritage. In 1989, a follow-up survey reexamined 181 of the 217 women who participated in the baseline survey. Nonrespondents to the follow-up survey were less likely to be married, were younger, had a greater Quetelet index, and had a slightly greater BMD than participants at baseline	Subjects were interviewed at baseline and follow-up for various factors, including lactation.	

Sowers, 1992 [UI#4543]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments																																				
BMD was expressed as the bone mineral to bone width ratio. Percent change was calculated as 100% - follow-up value/baseline value	Multiple-variable regression analyses were used to assess relationships between BMD change or baseline BMD values and the explanatory variables adjusting for age and estrogen status. Stepwise regression was used to generate models considering all other important variables, including age, estrogen status, parity, age at first pregnancy, and weight.	Nulliparous women (n=12) had significantly lower BMD at baseline and follow-up than parous women (n=169) before and after adjusting for age and estrogen status. A recalled history of breast-feeding in parous women did not predict significant differences in BMD level or amount of BMD change. We had insufficient sample size to test whether duration of breast-feeding in excess of 6 months for a single infant or 6 months cumulative time from breast-feeding multiple infants was associated with significant differences in BMD.	<table border="1"> <tr> <td>A: strong, B: moderate, C: weak</td><td>A</td><td>B</td><td>C</td></tr> <tr> <td>Selection</td><td></td><td>X</td><td></td></tr> <tr> <td>Study design</td><td></td><td>X</td><td></td></tr> <tr> <td>Confounder</td><td></td><td>X</td><td></td></tr> <tr> <td>Blinding</td><td></td><td>X</td><td></td></tr> <tr> <td>Data collection</td><td></td><td>X</td><td></td></tr> <tr> <td>Withdraw and dropout</td><td></td><td></td><td>X</td></tr> <tr> <td>Analyses</td><td></td><td>X</td><td></td></tr> <tr> <td>Intervention integrity</td><td></td><td></td><td></td></tr> </table> Overall: B	A: strong, B: moderate, C: weak	A	B	C	Selection		X		Study design		X		Confounder		X		Blinding		X		Data collection		X		Withdraw and dropout			X	Analyses		X		Intervention integrity			
A: strong, B: moderate, C: weak	A	B	C																																				
Selection		X																																					
Study design		X																																					
Confounder		X																																					
Blinding		X																																					
Data collection		X																																					
Withdraw and dropout			X																																				
Analyses		X																																					
Intervention integrity																																							

Uusi-Rasi, 2002 [UI#11991440]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 28.4 Menopause: pre Race: ND Enrolled/Evaluate: 132/92 Location: Finland Sites: Single/Multi Funding: private	healthy	Prospective cohort 4.2 (range 3.2-5.4) years	Healthy, non-smoking women willing to participate in original cross-sectional study. Inclusion criteria were a daily dietary calcium intake of over 1200 mg (Ca+) or under 800 mg (Ca-) and vigorous physical activity more than twice a week, or light to moderate activity no more than once a week. Women who had intermediate levels of physical activity or calcium intake or used calcium supplements were excluded. Altogether 92 women out of the 132 original 25- to 30-year-old subjects participated in the follow-up measurements.	Total duration of breastfeeding (one long period or sum of many periods of breastfeeding)	

Uusi-Rasi, 2002 [UI#11991440]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
BMC at the proximal femoral neck and total trochanter region and the distal radius of the dominant side was measured with dual energy X-ray absorptiometry.	Repeated-measures analyses of variance were used to analyze the between group differences in changes in site-specific BMC. Regression models for % bone loss using a multivariate stepwise regression analysis that included baseline variables: age, body weight, muscle strength, estimated maximal oxygen uptake and calcium intake, the absolute changes in body weight, muscle strength, estimated maximal oxygen uptake, calcium intake, physical activity, duration of breastfeeding and time from weaning.	The mean decrease in BMC (95% CI) was 1.5% (0.7% to 2.4%) for the femoral neck, 0.6% (-0.8% to 1.9%) for the trochanter and 6.0% (4.5% to 7.4%) for the distal radius during the follow-up According to the multiple stepwise regression analyses, the statistically significant independent predictors for site-specific bone loss were low calcium intake at the baseline and change in body weight both at the proximal femur and at the distal radius. High calcium intake, as well as an increase in body weight, were associated with less bone loss at all the measured bone site. In addition, breastfeeding was associated with radial bone loss; the longer the duration of breastfeeding the greater the bone loss (Multiple regression coefficient= -0.34 +/- 0.14 SE, p=0.015). Breastfeeding was not associated with femoral neck or trochanter BMC loss	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder			x
			Blinding		x	
			Data collection		x	
			Withdraw and dropout			x
			Analyses		x	
			Intervention integrity			
			Overall: B			

Alderman, 1986 [UI#3728442]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 50-74 Menopause: post (assumed) Race: White Enrolled/Evaluate: 355/344 for cases; 562/318 for controls Location: US Sites: Multi Funding: Government	ND	Case-control study	Cases: Between 1976 and 1980, all women with hip fracture who sought care from a 62% sample of orthopedic surgeons were eligible. Women with forearm fracture who sought care from these orthopedists were eligible if the physician's practice had records that allowed tracing of all outpatients. Women who sought care for vertebral fractures were not eligible because we believed that they represented a small and probably unrepresentative group of women with vertebral fracture. Exclusion: nonwhite, younger than 50 or older than 74 years of age, nonresidents of King County, residents of institutions at the time of fracture, or if they had multiple or pathologic fractures. Controls: unmatched control women were identified through 3 door-to-door area-based surveys of the general population of King Country conducted in 1976, 1977, and 1979. Comparable exclusion criteria were used.	The duration of lactation was calculated by summing the number of months of lactation associated with each birth, and was treated in separate analyses as a categorical variable (taking values 0, 1 to 12, 13 to 24 and more than 24 months) and as continuous variable (taking values from 0 to 109 months) 11/355 cases missing data on lactation	253/562 control women missing data on lactation

Alderman, 1986 [UI#3728442]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Risk of hip or forearm fractures	Logistic regression was used. The logistic model was evaluated for confounding by forward selection method. Only parity and variables with independent confounding effects were included in the final model. The same procedure was repeated to examine the association between fracture and duration of lactation.	Logistic regression analysis did not reveal a substantial change in risk of fracture with increasing duration of lactation when lactation was treated as a categorical variable. Adjustment for confounding effects of age, relative weight, and duration of estrogen use did not alter these results. Similar results were obtained when lactation was treated as a continuous variable. Separate analysis of women with hip fracture suggested that an increase in months of breastfeeding was associated with an irregular decreased in risk of fracture.	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder		x	
			Blinding			x
			Data collection			x
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity			
			Overall: B			

Chan, 1996 [UI#8783297]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 75 (70-79) Menopause: post Race: Chinese Enrolled/Evaluate: 481/481 Location: Hong Kong Sites: Single Funding: Government	ND	Case-control study	Chinese women aged 70-79 years who were living in Geriatric Housing Scheme in Hong Kong. Subjects with a known history of metabolic bone disorder, cancer, or who were on long-term steroid treatment, were excluded. No subjects with such histories but 3 subjects on steroid therapy were excluded from the study. Cases: definite or doubtful vertebral fracture Control: the remaining enrolled subjects (women with no fractures)	Home-visit interviews by a standardized, structured questionnaire was administered by 2 trained research nurses who was blinded to the fracture status. Period of lactation: less than 24 months, or 24 months or more	Period of lactation: never

Chan, 1996 [UI#8783297]

Outcome Definition	Statistical analyses and confounders adjusted	Results						Bias/limitations Comments			
Diagnosis of vertebral fracture according to Black et al.: definite vertebral fracture: when any of the 3 vertebral ratios was 3 SD or more away from the mean; doubtful fracture: any of the 3 vertebral ratios was 2-2.99 SD away from the mean. A woman with 1 or more definite fractures was classified as a definite case, a woman with no definite fractures but one or more doubtful fractures as a doubtful case, and the rest as controls.	Logistic regression: OR for vertebra fracture among definite case and control; doubtful cases and controls. Analysis adjusted for age.	There were no significant differences in the age of menarche and menopause between definite cases, doubtful cases, and controls. The risk of definite fracture decreased with the number of children given birth to and the duration of breast-feeding.						A: strong, B: moderate, C: weak	A	B	C
		Period of lactation	% in definite cases (n=144)	% in doubtful cases (n=174)	% in controls (n=163)	Age-adj OR (95% CI)		Selection		X	
						Definite cases & control	Doubtful cases & control	Study design		X	
								Confounder			X
								Blinding	X		
								Data collection		X	
								Withdraw and dropout		X	
								Analyses		X	
								Intervention integrity	N/A		
						Overall: C Did not separate parous women.					

Cumming, 1993 [UI#8225744]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 65-100 Menopause: post Race: D Enrolled/Evaluate: Location: Australia Sites: Multi Funding: ND	Self-reported health (for directly interviewed subjects only: 104 cases and 114 controls) Cases: 23% excellent, 50% good, 20% fair; 6% poor Controls: 26% excellent, 51% good, 21% fair; 5% poor	Population-based case-control study	All people aged ≥ 65 living in a postcode-defined area in Sydney between 3/1990 and 8/1991. Total of 11 hospitals were included in the study. Cases: People with hip fractures presenting to the hospitals. The case ascertainment varied across hospitals. Controls: People living in private homes in the community and from people living in nursing homes and hostels. The controls comprised a probability sample of people aged ≥ 65 years living in the 11 postcodes of the study base during the same time period as the cases. Proxy respondents were used for elderly people with cognitive impairment and those with various other health problems. A shortened questionnaire was used to interview proxy respondents.	For all women: Ever breastfeeding For parous women only: Average duration of breastfeeding per child (months)	For all women: Never breastfeeding

Cumming, 1993 [UI#8225744]

Outcome Definition	Statistical analyses and confounders adjusted	Results					Bias/limitations Comments				
Risk of hip fracture	Data analysis: crude, stratified and multivariate methods to produce OR for associations between hip fracture and exposure variables. Age is the strongest predictor of hip fracture and so all results were adjusted for age. Multiple logistic regressions were controlling for several confounding variables. Selection of confounders for inclusion in these models was based a priori knowledge of risk factors for hip fracture and the strength of univariate associations between exposure variables and hip fracture. Confounders (other than age) were entered into models as categorical variables.	There were 311 women in this study: 174 cases and 137 controls. The response rate was 96% for cases and 82% for controls. Proxy interviews were required for 93 subjects. Most cases (78%) were recruited from 1 teaching hospital. A history of use of HRT was more common among controls than among cases (age-adj OR: 0.55, 95%CI 0.24-1.27).					A: strong, B: moderate, C: weak	A	B	C	
							Selection		X		
							Study design		X		
							Confounder		X		
							Blinding			X	
							Data collection		X		
							Withdraw and dropout			X	
							Analyses		X		
							Intervention integrity	N/A			
							Overall: B				
							Missing data for covariates in the multivariate model, so the number of subjects in the model is limited.				
					</						

Hoffman, 1993 [UI#8338971]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 50-103 Menopause: post Race: White Enrolled/Evaluate: 286/174 cases; 311/174 controls Location: USA Sites: Multi Funding: ND	Hospital cases and controls	Case-control study	Cases: a sample of radiologically confirmed cases of first hip fracture among white women aged 45 years and older was drawn from among admissions to one of 30 hospitals between 9/1987 and 7/1989 Controls: white hospital controls, frequency-matched to cases by 10-year age group and by hospital of admission during the same time period. Exclude subjects with a prior hip fracture, pathological fracture secondary to cancer, evidence of bony metastases, severe cognitive, language, or hearing impairment, died in the hospital shortly after surgery, or medically unstable	Ever lactated Lactated ≤ 12 mo Lactated > 12 mo	Never lactated

Hoffman, 1993 [UI#8338971]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Risk of hip fracture	Conditional logistic regression models were used to remove the effect of matching cases and controls. Because the age intervals were wide, age was controlled for as continuous variable. Variables that changed crude estimates by 10% or greater were retained as confounding variables. For lactation, only the number of live births was retained.	Lactation was associated with reduced odds of hip fracture among all women (OR=0.66 [0.41-1.05]) Restricting the analysis to parous women and controlling for number of live births, there was no association between ever having lactated or duration of lactation and hip fracture (OR for lactation of 12 months or less=0.80 [0.42-1.55]); OR for lactation of more than 12 months=1.08 [0.45-2.60]) Duration of lactation did not confound the negative association between parity and hip fracture.	A: strong, B: moderate, C: weak	A	B	C
			Selection			X
			Study design		X	
			Confounder		X	
			Blinding			X
			Data collection		X	
			Withdraw and dropout			X
			Analyses		X	
			Intervention integrity			
			Overall: B			

Kreiger, 1982 [UI#7102649]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 45-74 Menopause: post Race: ND Enrolled/Evaluate: 149/98 for cases; 1345/884 Location: US Sites: Multi Funding: ND	Hospital cases and controls	Case-control study	Women aged 45-74 years admitted to surgical services of 4 Connecticut hospitals between 9/1977 and 5/1979. Cases: women admitted to 1 of the 4 study hospitals with a first diagnosis of hip fracture, either intra- or extra- capsular, confirmed by x-ray. Controls: systematic sample of women admitted to the inpatient surgical services of the 3 largest hospitals. Controls included had a wide variety of diagnoses, and no more than 5% of the controls were in any single diagnostic category. Exclusion: women with reproductive cancers, who did not speak English, who lived outside of Connecticut, or last menstrual period was within 1 year of hospital admission or if the admitting diagnosis of a disease related to estrogen use.	98 of 149 cases (66%) of hip fracture and 884 of 1345 controls (66%) were interviewed. Variable: Breastfeeding 12-month increase	Before the data analysis began, the controls were divided into 2 groups: trauma control group and nontrauma control.

Kreiger, 1982 [UI#7102649]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments																																														
Odds ratio of hip fracture, compared cases to trauma or nontrauma controls.	Linear logistic regression, adjusted for age and for Quetelet index (similar to BMI), bilateral ovariectomy, and estrogen replacement therapy.	Linear logistic regression, adjusted for age and for Quetelet index (0.01 increase), bilateral ovariectomy (yes/no), and estrogen replacement therapy (60-month increase): <table> <tr> <th rowspan="2">Variable</th><th rowspan="2">Unit</th><th>Cases and trauma control</th><th>Cases and non-trauma control</th></tr> <tr> <th>Adj OR (95% CI)</th><th>Adj OR (95% CI)</th></tr> <tr> <td>Breastfeeding</td><td>12-month increase</td><td>0.5 (0.2, 0.9)</td><td>0.6 (0.3, 1.0)</td></tr> </table>	Variable	Unit	Cases and trauma control	Cases and non-trauma control	Adj OR (95% CI)	Adj OR (95% CI)	Breastfeeding	12-month increase	0.5 (0.2, 0.9)	0.6 (0.3, 1.0)	<table> <tr> <td>A: strong, B: moderate, C: weak</td><td>A</td><td>B</td><td>C</td></tr> <tr> <td>Selection</td><td></td><td>x</td><td></td></tr> <tr> <td>Study design</td><td></td><td>x</td><td></td></tr> <tr> <td>Confounder</td><td></td><td>x</td><td></td></tr> <tr> <td>Blinding</td><td></td><td></td><td>x</td></tr> <tr> <td>Data collection</td><td></td><td></td><td>x</td></tr> <tr> <td>Withdraw and dropout</td><td></td><td></td><td>x</td></tr> <tr> <td>Analyses</td><td></td><td></td><td>x</td></tr> <tr> <td>Intervention integrity</td><td></td><td></td><td></td></tr> </table> Overall: C Sample size in breastfeeding analysis was not described. Did not separate out parous women.	A: strong, B: moderate, C: weak	A	B	C	Selection		x		Study design		x		Confounder		x		Blinding			x	Data collection			x	Withdraw and dropout			x	Analyses			x	Intervention integrity			
Variable	Unit	Cases and trauma control			Cases and non-trauma control																																												
		Adj OR (95% CI)	Adj OR (95% CI)																																														
Breastfeeding	12-month increase	0.5 (0.2, 0.9)	0.6 (0.3, 1.0)																																														
A: strong, B: moderate, C: weak	A	B	C																																														
Selection		x																																															
Study design		x																																															
Confounder		x																																															
Blinding			x																																														
Data collection			x																																														
Withdraw and dropout			x																																														
Analyses			x																																														
Intervention integrity																																																	

Michaelsson, 2001 [UI#2001082617]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Cases=72.5; Controls=70.5 Menopause: post Race: ND Enrolled/Evaluate: 715/664 for cases evaluated Location: Sweden Sites: Multi Funding: ND	ND	Case-control study; a comprehensive questionnaire at a mean interval of 95 days after the fracture; associations between lactation and hip fracture were only analyzed for parous women	Cases: All fractures of the proximal femur that occurred between 10/ 1993 and 2/1995 among women resident in 6 counties in Sweden who were born after 1913. Using clinical records or operation registers in all 24 hospitals in the study area, a total of 1,597 possible incident cases were identified. Those with a fracture due to malignancy (n=26); high-energy trauma (n=4); incorrect diagnosis (n=41); old fracture (n=10); blindness (n=5); birth outside Sweden (n=202); excluded: severe alcoholic abuse, psychosis, or senile dementia (n=576); or death within 3 months of the fracture (n=123). After exclusions, 1,610 eligible cases remained and were approached with. The Swedish inpatient register identified 34 additional cases. Controls: Native-born women were randomly selected from the national population register in the month before the start of the study. Potential controls aged 70-80 years were frequency matched (2 controls per 1 case) to the expected hip fracture age distribution within county of residence. Controls aged 50-69 years were randomly selected from the population register as part of a breast cancer study being conducted at the same time with the same questionnaire. For these women, frequency matching to the expected number of breast cancer cases provided 2-4 times as many controls as hip fracture cases in each 5-year age group and country of residence. Of the 4,870 candidate controls in the hip fracture analysis, 4,059 were eligible for the study, 610 were born outside of Sweden, 257 died before being approached, 44 were senile or psychotic, and 2 were blind.	Duration of breastfeeding was considered in 4 classes defined by the quartiles of either total duration or mean duration per child of breastfeeding among controls	

Michaelsson, 2001 [UI#2001082617]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
OR of hip fracture	Unconditional logistic regression were used as measure of association. The multivariate analyses included the covariates age, BMI, oral contraceptive use, HRT use, and smoking status. Interactions were considered. Associations between lactation and hip fracture were only analyzed for parous women	Long total duration of breastfeeding was associated with a reduction in risk (table 3 in paper), but this association disappeared after adjustment for parity, body mass index, and use of exogenous estrogens. There were no substantial risk differences in analyses that considered mean duration of breastfeeding per child (quartiles 0–2, 3–4, 5–6, and 7 months per child or more) (data not shown). Long duration of breastfeeding also had no substantial association with hip fracture risk among those with their first pregnancy as a teenager or among those with their first pregnancy after age 30 years. The relative risk estimates for lactation and parity were similar for cervical and trochanteric hip fracture (data not shown).	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design		x	
			Confounder		x	
			Blinding			x
			Data collection		x	
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity			
			Overall: B			

Appendix C. Evidence Tables

Part III. Maternal Outcomes

Ovarian Cancer

Chiaffarino 2005
Cramer 1983
Greggi 2000
Gwinn 1990
Hartge 1989
John 1993
Modugno 2003
Ness 2000
Riman 2002
Risch 1983
Risch 1994
Siskind 1997
Titus-ernstoff 2001
Tung 2003, 2005
West 1966
Whittemore 1992
Wynder 1969
Yen 2003

Chiaffarino, 2005 [UI# 15975644]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Median age (range): Cases: 56 (18-79) Controls: 57 median (17-79) Menopause: ND Race: ND Enrolled/Evaluate: Cases: 1031/1028 Controls: 2411/2390 Location: Italy Sites: Multi-center Funding: Private, government	Cases: ND Controls: traumatic conditions 26%, nontraumatic orthopedic disorders 28%, acute surgical conditions mostly abdominal diseases 15%, miscellaneous illnesses (ear, nose throat, teeth) 31%	Case-control Data collection by interviews using structured questionnaire	Case: Histologically confirmed epithelial ovarian cancer diagnosed in past year, admitted to major or teaching hospital Control: residents of same geographic area, admitted to same network of hospitals for acute, non-neoplastic, unrelated to known or likely risk factors for ovarian cancer, exclusion include hormonal and gynecological diseases, or ovariectomized	Breast feeding status determined by questionnaire	Breast feeding status determined by questionnaire

Chiaffarino, 2005 [UI# 15975644]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Ovarian cancer confirmed by histology	Unconditioned multiple logistic regression adjusting for age, center, education, parity, oral contraceptive use, family history of ovarian/breast cancer in first degree relatives	Exposure status	Cases n (%)	Controls n (%)	OR _{adj} (95% CI)	A: strong, B: moderate, C: weak	A	B	C
		Never breastfed	368(36)	878 (37)	1.0				
		Ever breastfed	660 (64)	1512 (63)	1.16 (0.93-1.43)	Selection		X	
		Breastfed 1-4 months	157 (15)	358 (15)	1.20 (0.91-1.59)	Study design		X	
		Breastfed 5-8 months	180 (18)	366 (15)	1.24 (0.95-1.62)	Confounder			X
		Breastfed 9-16 months	195 (19)	451 (19)	1.01 (0.77-1.33)	Blinding		X	
		Breastfed ≥17 months	128 (12)	337 (14)	1.21 (0.85-1.71)	Data collection			X
		x ² for trend			0.03 (p=0.87)	Withdraw and dropout		X	
		Odds ratio _{adj} (95% CI)				Analyses	X		
		Exposure status				Intervention integrity			
			Serous	Mucinous	Other subtypes	Potential recall bias			
		Never breastfed	1	1	1				
		Ever breastfed	1.1 (0.85-1.48)	1.59 (0.82-3.07)	1.08 (0.71-1.62)				
		Breastfed 1-4 mo	1.3 (0.90-1.85)						
		Breastfed 5-8 mo	1.16 (0.81-1.65)						
		Breastfed 9-16 mo	1.06 (0.74-1.51)						
		Breastfed ≥17 mo	0.87 (0.55-1.39)						
		x ² for trend	0.14 p=0.71						

Cramer, 1983 [UI# 6578366]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age: 53 Menopause: ND Race: White Enrolled/Evaluate: Cases: 215 Controls: 215 Location: USA Sites: Multi-center Funding: Government	ND	Case-control Data collection by questionnaires	Cases: white females with epithelial ovarian cancer residing in greater Boston area Controls: matched residents of same precinct, age within 2 years, race. Only bilateral oophorectomy excluded.	Breastfeeding status determined by questionnaire	Breastfeeding status determined by questionnaire

Cramer, 1983 [UI# 6578366]

Cramer, 1985 [61/ 65/6566]						
Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
All types of epithelial ovarian cancer	Chi-square	Crude relative risks 1.11 (0.70-1.76) never breastfed parous cases (86/137) vs controls (106/176) Non-significant trend in risk for ovarian cancer with number of children breastfed.	A: strong, B: moderate, C: weak	A	B	C
			Selection		X	
			Study design		X	
			Confounder			X
			Blinding		X	
			Data collection			X
			Withdraw and dropout		X	
			Analyses			X
			Intervention integrity			
			Minimal reporting of data Potential recall bias			

Greggi, 2000 [UI# 11006030]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Median age (range): Cases: 54 (13-80) Controls: 55 (19-80) Menopause: mix Race: ND Enrolled/Evaluate*: Cases: 440/330 Controls 868/721 Location: Italy Sites: Single Funding: Private *Parous	Cases: ND Controls: traumatic conditions 29%, nontraumatic orthopedic disorders 21%, acute abdominal diseases generally requiring surgery 17%, miscellaneous illnesses (ear, nose throat, teeth) 33%	Case-control Data collection by hospital interviews	Cases: admitted histologically confirmed diagnosis of epithelial ovarian cancer Controls: admitted to same hospital, similar age strata, acute nongynecological, nonhormonal, or nonneoplastic condition	Breast feeding status determined by interview	Breast feeding status determined by interview

Greggi, 2000 [UI# 11006030]

Outcome Definition	Statistical analyses and confounders adjusted	Results					Bias/limitations Comments			
Epithelial ovarian cancer confirmed by histology	Mantel-Haenzel adjusted for age; unconditional multiple logistic regression estimates including terms for age, education, parity, oral contraceptive use, family history of ovarian cancer	Exposure status*	Cases 330 (%)	Controls 721 (%)	OR _{adj} (95% CI) Mantel-Haenzel Multivariate regression		A: strong, B: moderate, C: weak	A	B	C
		Never breastfed	84 (25)	120 (17)	1.0	1.0	Selection		X	
		Breastfed ≤12 months	136 (41)	294 (41)	0.6 (0.5-0.9)	0.8 (1.0-1.1)	Study design		X	
		Breastfed >12 months	110 (33)	307 (43)	0.5 (0.3-0.7)	0.5 (0.4-0.8)	Confounder		X	
		*Parous					Blinding		X	
							Data collection			X
							Withdraw and dropout			X
							Analyses		X	
							Intervention integrity			
		Potential recall bias								

Gwinn, 1990 [UI# 2348208]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): ND Menopause: mix Race: ND Enrolled/Evaluate: Cases: 494/436 Controls: 4236/3833 Location: USA Sites: Multi center Funding:	ND	Case-control Controls contacted by random telephone Data collection by interview of standard questionnaire	Cases: 20-54 year, resides in 1 of 8 study areas, ovarian cancer by histology Controls: matched for age, resides in same 8 areas, no ovaries or unknown number of ovaries are exclusion	Breast feeding status determined by interview, duration in months regardless of pattern	Breast feeding status determined by interview

Gwinn, 1990 [UI# 2348208]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Epithelial ovarian cancer confirmed by histology	Logistic regression adjusted for pregnancy, oral contraceptive use, age, and age-pregnancy interaction	Breastfeeding duration (month)*	Cases n (%)	Controls n (%)	RR _{adj} (95% CI)	A: strong, B: moderate, C: weak	A	B	C
		0	184 (57)	1517 (46)	1.0	Selection		X	
		1-2	40 (12)	552 (17)	0.6 (0.5-0.9)	Study design		X	
		3-5	32 (10)	330 (10)	0.8	Confounder		X	
		6-11	33 (10)	379 (11)	0.8	Blinding		X	
		12-23	25 (8)	321 (10)	0.7	Data collection		X	
		≥ 24	7 (2.2)	213 (6.4)	0.3	Withdraw and dropout		X	
		* Parous women only				Analyses	X		
		β -0.024				Intervention integrity			
		Each month of breastfeeding reduced risk by 2.4%				Ovarian cancer confirmed by histology in majority of cases			
		Most protection due to first exposure rather than number of pregnancies							

Hartge, 1989 [UI# 2750791]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Cases: 54 (20-79) Controls: 55 Menopause: mix Race: Cases: 12% black Controls: 14% black Enrolled/Evaluate: Cases: 296/203* Controls: 343/257* Location: USA Sites: Multi Funding: Government *Parous women (evaluated)	Serous 28% Serous, low malignant potential 12% Mucinous 6% Mucinous, low malignant potential 3% Endometrioid, low malignant potential 26% Clear cell 4% Mixed epithelial 8% Undifferentiated 10% Controls: Infectious 2% Neoplasm 5% Endocrine-metabolic 5% Blood disease or blood-forming organs 1% Nervous system 8% Eye, ear, mastoid 8% Varicose veins, hemorrhoids 1% Respiratory 8% Digestive system 16% Urinary 5% Skin 2% Musculoskeletal 22% Congenital 1% Ill-defined 5% Fractures/injuries 11%	Case-control Data collection with interviewer of standard questionnaire and medical records review	Cases: primary epithelial ovarian cancer confirmed by microscopic slides and medical records, included low malignant potential tumors as well as malignant tumors Controls: hospitalized for other conditions, matched for age, hospital, race, at least one ovary. Excluded if psychiatric diagnosis or related to major exposures of interest	Breastfeeding status determined by interview	Breastfeeding status determined by interview

Hartge, 1989 [UI# 2750791]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Primary epithelial ovarian cancer confirmed by microscopic slides and medical records	Rate ratio of ovarian cancer incidence in exposed group to corresponding unexposed group, adjusted for confounding variables by stratified contingency table analysis and by logistic regression models, adjustments for age and race, and parity, difficulty conceiving, oral contraceptives, surgical menopause, HRT, or family history of ovarian cancer as needed	Breastfeeding duration (month)*	Cases n (%)	Controls n (%)	Rate ratio _{adj} (95% CI)	A: strong, B: moderate, C: weak	A	B	C
		0	112 (55)	121 (47)	1.0	Selection		X	
		1-9	62 (31)	84 (33)	0.8 (0.5-1.2)	Study design		X	
		10-18	16 (7.9)	37 (18)	0.5 (0.2-0.9)	Confounder		X	
		19-110	13 (6.4)	15 (7.4)	1.1 (0.5-2.6)	Blinding		X	
		* Parous women only Adjusted for age and race Trend test p = 0.14				Data collection		X	
						Withdraw and dropout		X	
						Analyses		X	
						Intervention integrity			
		Number of months breastfeeding not related to risk							

John, 1993 [UI# 8418303]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Cases: Invasive 53 Borderline 37 Controls: ND Menopause: mix Race: Black Enrolled/Evaluate: Cases: 110/53 Controls: 365/259 69% population 31% hospital Location: USA Sites: 7 of 12 studies Funding: Government	Invasive cancer 65% Low malignant 32% Unknown 3%	Case-control Data from 7 studies of cases diagnosed between 1971 and 1986. Individual patient data collected by interview from 7 case-control studies conducted in US	Epithelial ovarian cancer Controls: population and hospital subjects, bilateral oophorectomy excluded	Breastfeeding status determined by interview	Breastfeeding status determined by interview

John, 1993 [UI# 8418303]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Epithelial ovarian cancer	Conditional logistic regression adjusted for study, birth year, reference age, parity	Breastfeeding duration (month)*	Cases n (%)	Controls n (%)	RR _{adj} (95% CI)	A: strong, B: moderate, C: weak	A	B	C
		Never breastfed	24 (45)	102 (39)	1.0	Selection		X	
		Ever breastfed	29 (55)	157 (61)	0.90 (0.42-1.9)	Study design		X	
						Confounder			X
		1-5	11 (21)	52 (20)	1.0 (0.39-2.6)	Blinding		X	
		≥6	18 (34)	103 (40)	0.85 (0.36-2.0)	Data collection			X
		Trend per month			0.99 p=0.57	Withdraw and dropout		X	
						Analyses		X	
						Intervention integrity			
		Data unavailable for use of oral contraceptives and years of ovulation							

Overlap with Hartge 1989 and Cramer 1983

Modugno, 2003 [UI# 12946038]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): BRCA1 51 BRCA2 61 BRCA- 58 Menopause: mix Race: 100% Jewish Enrolled/Evaluate: Cases: 95 Controls: 147 Location: USA, Israel Sites: Multi Funding: Government	ND	Case-control Cases screened for Ashkenazi founder mutations BRCA1 (185delAG or 5382insC) or BRCA2 (6174delT) Controls were ovarian cases without BRCA1/2 genotype Individual patient data collected by interview from 4 case-control studies. Two US population studies (100 cases), one hospital-based study from 11 US and Israel centers (208 cases), and one US genetic counseling center (14 cases)	Jewish women with epithelial ovarian cancer, no breast cancer history, BRCA1/2 confirmed by data sequencing	ND	ND

Modugno, 2003 [UI# 12946038]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments				
Invasive epithelial ovarian cancer confirmed by histology	Unconditioned logistic regression adjusted for age at diagnosis, year of birth, number of live births, oral contraceptive use, history of tubal ligation. Mantel-Haenszel test for heterogeneity		OR _{adj} (95% CI)				A: strong, B: moderate, C: weak	A	B	C
			BRCA- (n=147)	BRCA1/2 (n=95)	BRCA1 (n=64)	BRCA2 (n=31)	Selection			X
		Breastfeeding	1.0	1.09 (0.61-1.97)	1.36 (0.68-2.73)	0.70 (0.28-1.72)	Study design		X	
		Breastfeeding duration*	1.0	1.02 (0.99-1.04)	1.01 (0.98-1.04)	1.02 (0.99-1.05)	Confounder			X
		* Parous women				Blinding		X		
						Data collection			X	
						Withdraw and dropout		X		
						Analyses		X		
						Intervention integrity				
							10% overlap in samples between the US genetic center & hospital-based study Classification of Jewish varied between studies			

One population study, Ness 2000, has been reported

Ness, 2000 [UI# 11021606]

Modugno, 2001 [UI# 11308435]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Cases: Invasive 53 Borderline 45 Controls: 49 Menopause: mix Race: ND Enrolled/Evaluate: Cases: 767/531* Controls: 1215/1190* Location: USA Sites: Multi Funding: Government *Parous women (evaluated)	Invasive epithelial 616 Serous 45% Mucinous 8% Endometrioid 22% Other 24% Borderline epithelial 151 Serous 52% Mucinous 40% Endometrioid 2% Other 6% Other includes clear cell, mixed cell, undifferentiated or poorly differentiated	Case-control Data collection by interview Controls identified by random-digit dialing from HFCA lists	Histologically confirmed epithelial ovarian cancer diagnosed past 6 months, 20-69 years Controls: ≤ 65 years, matched by 5 years and 3 digit exchange Exclusion: outside counties of referral hospital, prior diagnosis of ovarian cancer, non-English speaking, mentally incompetent, critically ill, prior bilateral oophorectomy	Breastfeeding status determined by interview, length recorded on life calendar	Breastfeeding status determined by interview, length recorded on life calendar

Ness, 2000 [UI# 11021606]

Modugno, 2001 [UI# 11308435]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Epithelial ovarian cancer confirmed by histology	Multivariate unconditional logistic regression adjusted for age, number of pregnancies, family history of ovarian cancer, race, oral contraceptive use, tubal ligation, hysterectomy, breastfeeding For analyses by histology: multivariate unconditional logistic regression adjusted for age, number of live births, years of oral contraceptive use, years of noncontraceptive estrogen use and months breastfed, tubal ligation, hysterectomy, family history of ovarian and breast cancer, ethnicity	Breastfeeding duration (month)*	Cases n (%)	Controls n (%)	Odds ratio _{adj} (95% CI)	A: strong, B: moderate, C: weak	A	B	C
		0	299 (56)	577 (48)	1.0	Selection		X	
		1-5	117 (22)	259 (22)	0.9 (0.7-1.2)	Study design		X	
		6-11	46 (8.7)	119 (10)	0.9 (0.6-1.3)	Confounder		X	
		12-23	40 (7.5)	124 (10)	0.7 (0.5-1.1)	Blinding			X
		≥24	29 (5.5)	111 (9.3)	0.6 (0.4-1.0)	Data collection		X	
		* Parous women only				Withdraw and dropout			X
		Histologic subgroup (N)	Months breastfed Odds ratio _{adj} (95% CI)			Analyses		X	
		Serous (357)	1.00 (0.98-1.01)			Intervention integrity			
		Mucinous (112)	1.00 (0.98-1.02)						
		Endometrioid (139)	0.97 (0.94-1.00)						
		Other (159)	0.97 (0.94-1.00)						
		All (767)	0.99 (0.98-1.00)						

Riman, 2002 [UI# 12181107]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Cases: 62 Controls: 63 Menopause: mix Race: ND Enrolled/Evaluate: Cases: 914/655 Controls: 4148/3899 Location: Sweden Sites:NA Funding: Private, government	All invasive (655) Serous (337) Mucinous (60) Endometrioid (180) Clear cell (43) Other/undifferentiated (35)	Case-control Data collected by self-administered questionnaires followed up by phone for missing or inconsistent details	Swedish born women 50-74 years old with epithelial ovarian cancer confirmed by histology identified by national registry Controls: randomly selected population national registry Exclusion: previous ovarian malignancy or bilateral oophorectomy	Breastfeeding status determined by questionnaires	Breastfeeding status determined by questionnaires

Riman, 2002 [UI# 12181107]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Epithelial ovarian cancer confirmed by histology	Unconditional logistic regression adjusted for age, bmi, parity, age at menopause, duration of oral contraceptive use, HRT use	Breastfeeding duration (months)	Cases n (%)	Controls n (%)	OR _{adj} (95% CI)	A: strong, B: moderate, C: weak	A	B	C
		<1	33 (7.2)	161 (15)	1.0	Selection	X		
		1-5	134 (29)	612 (15)	0.99 (0.64-1.52)	Study design		X	
		6-11	134 (29)	840 (19)	0.77 (0.50-1.19)	Confounder		X	
		≥12	158 (34)	1,024 (14)	0.87 (0.56-1.35)	Blinding	X		
		Parous women				Data collection		X	
		Breastfeeding duration (months)	Odds Ratio _{adj} (95% CI)				Withdraw and dropout		X
			Serous	Mucinous	Endometrioid	Clear cell	Analyses		X
		<1	1.0	1.0	1.0	1.0	Intervention integrity		
		1-5	0.87 (0.50-1.53)	2.19 (0.49-9.87)	1.05 (0.47-2.34)	0.54 (0.16-1.87)	50% of cases and controls were contacted by phone for missing or inconsistent data; unknown if breastfeeding status category <1 month includes no breastfeeding or data was unavailable		
		6-11	0.61 (0.35-1.09)	1.75 (0.39-7.87)	1.10 (0.50-2.46)	0.23 (0.06-0.88)			
		≥12	0.87 (0.49-1.54)	0.83 (0.49-1.54)	1.02 (0.44-2.37)	0.24 (0.06-0.97)			

Risch, 1983 [UI# 6681935]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Cases: 20-74 Controls: 21-75 Menopause: mix Race: white Enrolled/Evaluate: Cases: 284 Controls: 705 Location: USA Sites: Multi Funding: Government	ND	Case-control Data collection by interview Cases	Epithelial ovarian cancer diagnosed in past 18 months, reside in 6 specified counties of Washington and Utah and identified by population-based cancer reporting systems, age 34-74 for Washington cases, age 20-74 for Utah cases Controls: household survey conducted in same counties by standard geographic sampling methods or telephone selected from directories. Utah samples matched for age and Washington samples matched for age range. Exclusion included bilateral oophorectomy $\leq$ 1 year Additional case-control exclusion: nonwhite, incomplete reproductive histories	Breastfeeding status determined by interview	Breastfeeding status determined by interview

Risch, 1983 [UI# 6681935]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Epithelial ovarian cancer	Linear logistic regression adjusted for age at diagnosis 20-44, nulliparous, no miscarriages, < 1 year total exposure to combined oral contraceptives, nonobese	RR 0.79 per year of lactation p=0.034, is associated with decreased ovarian cancer risk Cases reported less time breastfeeding than controls 3+ total months lactation – RR 0.69 (0.50-0.96) p=0.026	A: strong, B: moderate, C: weak	A	B	C
			Selection			X
			Study design		X	
			Confounder		X	
			Blinding		X	
			Data collection			X
			Withdraw and dropout		X	
			Analyses		X	
			Intervention integrity			

Risch, 1994 [UI# 7942759]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Cases: 57 Controls: 58 Menopause: mix Race: 96% white Enrolled/Evaluate: Cases: 450 Controls: 564 Location: Canada Sites: Multi Funding: Government	ND	Case-control Data collection by interview using questionnaire incorporating life events calendar – conducted at subject's home	Primary malignant or borderline malignant epithelial ovarian cancer confirmed by histology, age 35-79, residents of Ontario located through Ontario Cancer Registry, living subjects Controls: obtained from Enumeration Composite Record listing compiled by Ontario Ministry of Revenue, matched for area of residence and age range, exclusion included bilateral oophorectomy < 1 year, living subjects	Breastfeeding status determined by interview	Breastfeeding status determined by interview

Risch, 1994 [UI# 7942759]

Outcome Definition	Statistical analyses and confounders adjusted	Results			Bias/limitations Comments				
Primary malignant or borderline malignant epithelial ovarian cancer confirmed by pathology reports	Multivariate unconditional continuous logistic regression adjusted for 3 age groups, continuous variables age, total duration of oral contraceptive use	No of full-term pregnancies	Mean		Odds ratio _{adj} (95% CI)	A: strong, B: moderate, C: weak	A	B	C
			Cases	Controls		Selection		X	
		Total duration* (year)	0.51	0.65	0.89 (0.75-1.05)	Study design		X	
		Average duration of lactation/pregnancy* (months)	2.24	2.72	0.87 (0.76-0.99)	Confounder			X
		*Adjusted for number of full-term pregnancies				Blinding		X	
		Inverse association with total duration of lactation				Data collection		X	
		Inverse association with average duration per pregnancy, p=0.030				Withdraw and dropout		X	
		Pregnancies with lactation slightly more protective than pregnancies without				Analyses			X
						Intervention integrity			
						Slightly greater number for controls born in Canada or US, smaller percentage of cases were parous or ever used OC and among those cases had fewer full-term pregnancies and used OC for shorter periods of time			

Siskind, 1997 [UI# 9229212]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age* (range): Cases: 57 Controls: 56 Menopause: mix Premenopausal Cases: 35% Controls: 36% Race: ND Enrolled/Evaluate*: Cases: 824/618 Controls: 855/724 Location: Australia Sites: Multi Funding: Government, private *Parous	ND	Case-control Data collection by interviewer with questionnaire at discharge or clinic follow-up, in homes for all controls and homes or clinics for cases Controls randomly selected through electoral roll	Cases: Histologically confirmed primary epithelial ovarian cancer, 18-79 years, ≥ 1 live birth Controls: similar age and region Exclusion: history of ovarian cancer or bilateral oophorectomy, incapable of completing questionnaire	Breastfeeding status determined by questionnaire	Breastfeeding status determined by questionnaire

Siskind, 1997 [UI# 9229212]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Primary epithelial ovarian cancer confirmed by histology	Multiple logistic regression with covariates: age, parity, age at 1 st birth, education, oral contraceptive use, smoking history, menopausal status	Breastfeeding duration (month)	Cases n (%)	Controls n (%)	Odds ratio _{adj} (95% CI)	A: strong, B: moderate, C: weak	A	B	C
		0	168 (27)	155 (21)	1.0	Selection		X	
		1-6	180 (29)	190 (26)	0.89 (0.65-1.2)	Study design		X	
		7-12	125 (20)	187 (26)	0.68 (0.49-0.94)	Confounder		X	
		13-24	107 (17)	141 (19)	0.84 (0.59-1.2)	Blinding			X
		25-36	25 (4)	39 (5)	0.69 (0.38-1.3)	Data collection		X	
		>36	13 (2)	12 (2)	0.77 (0.34-1.8)	Withdraw and dropout		X	
		Any	450 (73)	569 (79)	0.80 (0.61-1.04)	Analyses	X		
		Per month			0.99 (0.98-1.0)	Intervention integrity			
		* Parous women only							

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments										
Primary epithelial ovarian cancer confirmed by histology	Multiple logistic regression with covariates: age, parity, age at 1 st birth, education, oral contraceptive use, smoking history, menopausal status	Breastfeeding duration (month)	Premenopausal		Postmenopausal		A: strong, B: moderate, C: weak	A	B	C						
			Cases n (%)	Controls n (%)	Cases n (%)	Controls n (%)										
			Odds ratio _{adj} (95% CI)		Odds ratio _{adj} (95% CI)											
		0	75 (35)	66 (25)	93 (23)	89 (19)					Selection		X			
			1.00		1.00						Study design		X			
		1-6	69 (32)	78 (29)	111 (27)	112 (24)					Confounder		X			
			0.75 (0.46-1.21)		0.98 (0.65-1.47)						Blinding			X		
		7-12	34 (16)	72 (27)	91 (23)	115 (25)					Data collection		X			
			0.53 (0.31-0.94)		0.83 (0.54-1.26)						Withdraw and dropout		X			
		13-24	33 (15)	36 (14)	74 (18)	105 (23)					Analyses	X				
			1.03 (0.54-1.95)		0.88 (0.56-1.38)						Intervention integrity					
		>24	4 (4)	12 (5)	---											
			0.29 (0.08-1.04)													
		25-36	---		21 (5)	27 (6)										
					0.93 (0.46-1.88)											
		>36	---		13 (3)	12 (3)										
					1.27 (0.50-3.2)											
		* Parous women only														

Titus-Ernstoff, 2001 [UI# 11237375]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (%): Cases Controls <30 29(5.2) 37(7.1) 30-39 78(14) 89(17) 40-49 159(28) 136(26) 50-59 126(22) 118(23) 60-69 118(21) 117(22) ≥70 53(9.4) 26(5.0) Menopause: mix Race: ND Enrolled/Evaluate: Cases: 563/378* Controls: 523/417* Location: USA Sites: Multi Funding: Government * Parous women only	Serous borderline 86 Serous invasive 229 Mucinous 83 Endometrioid/clear cell 130 * Parous women only	Case-control Data collection by interviewer with questionnaire in homes for all controls and case Controls selected by random digit dialing, >60 years by Town Books	Cancer registry and hospital board cases of epithelial ovarian cancer including lesions of borderline malignancy, 20-74 years, Controls: matched by age within 4 years and telephone sampling unit, >60 years were matched by age and precinct, exclusion included bilateral oophorectomy	Breastfeeding status determined by interview	Breastfeeding status determined by interview

Titus-Ernstoff 2001 [UI# 11237375]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Epithelial ovarian cancer by pathology report or microscopy slides	Unconditional logistic regression adjusted for age, state, parity	Breastfeeding duration (month)*	Cases n (%)	Controls n (%)	Odds ratio _{adj} (95% CI)	A: strong, B: moderate, C: weak	A	B	C
		0	239 (63)	223 (53)	1.0				
		<3	43 (11)	48 (12)	0.9 (0.6-1.4)	Selection		X	
		3-6	54 (14)	76 (18)	0.7 (0.5-1.1)	Study design		X	
		>6	42 (11)	70 (17)	0.7 (0.4-1.0)	Confounder		X	
		Ever	139 (37)	194 (47)	0.7 (0.5-1.0)	Blinding			X
		P for trend			0.21	Data collection		X	
		* Parous women only				Withdraw and dropout		X	
		Breastfeeding Status	Odds Ratio _{adj} (95% CI)				Analyses		X
			Serous borderline (n=86)	Serous invasive (n=229)	Mucinous (n=83)	Endometrioid /Clear cell (n=130)	Intervention integrity		
		Never breastfed	1.0	1.0	1.0	1.0			
		Ever breastfed	0.8 (0.4-1.6)	1.0 (0.6-1.4)	0.6 (0.3-1.2)	0.4 (0.2-0.7)			
		P for trend ¹	0.61	0.34	0.88	0.04			
		¹ Average duration (months) per breastfed child							

Tung, 2003 [UI# 14507598], Tung 2005 [UI# 15692075]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions																				
Mean age (range): Cases: 50 Mucinous 55 Nonmucinous Controls: 56 Menopause: premenopausal 39% Race %: <table><tr><th></th><th>Cases</th><th>Cases</th><th></th></tr><tr><th></th><th>Mucinous</th><th>Nonmucinous</th><th>Controls</th></tr><tr><td>Asian</td><td>51</td><td>34</td><td>42</td></tr><tr><td>White</td><td>30</td><td>50</td><td>44</td></tr><tr><td>Other</td><td>19</td><td>16</td><td>14</td></tr></table> Enrolled/Evaluate: Cases: 558 Controls: 607 Location: USA Sites: Multi Funding: Government		Cases	Cases			Mucinous	Nonmucinous	Controls	Asian	51	34	42	White	30	50	44	Other	19	16	14	Invasive n=431 Mucinous 11% Serous 51% Endometrioid 17% Clear cell 11% Other(undifferentiated, squamous, transitional) 10% Borderline n=127 Mucinous 48% Serous 52%	Case-control Data collection by interviewer with questionnaire including life calendar in homes for 95% of all interviewed	Cases: ≥18 years, epithelial ovarian cancer confirmed by histology, identified by population-based cancer registry, reside in Los Angeles, California or Hawaii for 1 year prior to diagnosis Controls: ≥18 years, reside in Los Angeles, California or Hawaii for 1 year prior to Interview, no prior history of ovarian cancer, ≥1 intact ovary, matched for age (±5 years), ethnicity, study site, randomly selected from neighborhood walk procedure in LA and Hawaii Department of Health state-wide annual survey list	Breastfeeding status determined by interview	Breastfeeding status determined by interview
	Cases	Cases																							
	Mucinous	Nonmucinous	Controls																						
Asian	51	34	42																						
White	30	50	44																						
Other	19	16	14																						

Epithelial ovarian cancer confirmed by histology	Polytomous logistic model adjusted for age, ethnicity, study site, education, oral contraceptive use, tubal ligation Unconditional logistic regression adjusted for age, ethnicity, study site, education, tubal ligation, HRT, ovulation variables	Breastfeeding		Odds ratio _{adj} (95% CI)			
				All cases	Mucinous	Nonmucinous	
		Never		1	1	1	
		Ever		0.6 (0.4-0.7)	0.8 (0.5-1.4)	0.5 (0.4-0.7)	
		Breastfeeding duration (month)*					
		0		1	1	1	
		≤5		0.6 (0.4-0.9)	0.6 (0.4-1.4)	0.7 (0.4-0.9)	
		6-16		0.6 (0.4-0.9)	0.7 (0.3-1.3)	0.5 (0.3-0.7)	
		>16		0.6 (0.4-0.8)	0.9 (0.8-1.8)	0.4 (0.3-0.7)	
		P for trend*		0.011	0.99	0.0005	
*Based on likelihood ratio test comparing models with and without a trend variable that was assigned median values for the categories							
Total breastfeeding duration (year)*		Odds ratio _{adj} (95% CI)					
		Premenopausal		Postmenopausal			
0		1					
<0.5		0.96 (0.49-1.85)		0.64 (0.41-1.02)			
0.5-1.0		0.81 (0.43-1.54)		0.60 (0.36-1.00)			
>1.0		0.46 (0.22-0.97)		0.69 (0.43-1.12)			
P for trend		0.003		0.08			
Odds ratio _{adj} (95% CI)		Breastfeeding		Breastfeeding duration (month)*			
Invasive cases		Never	Ever	≤5	6-16	>16	P for trend*
All cases		1	0.5 (0.4-0.7)	0.6 (0.4-0.9)	0.5 (0.4-0.8)	0.5 (0.3-0.7)	0.002
Mucinous			1.2 (0.6-2.7)	0.7 (0.3-1.8)	1.1 (0.5-2.6)	1.2 (0.5-3.0)	0.30
Serous			0.5 (0.3-0.7)	0.6 (0.4-0.9)	0.5 (0.3-0.8)	0.4 (0.2-0.7)	0.96
Endometrioid			0.5 (0.3-0.9)	0.6 (0.3-1.2)	0.6 (0.3-1.3)	0.3 (0.1-1.0)	0.10
Clear cell			0.5 (0.2-1.0)	0.5 (0.3-1.5)	0.5 (0.3-1.4)	0.3 (0.1-1.1)	0.40
A: strong, B: moderate, C: weak							
Selection							
Study design							
Confounder							
Blinding							
Data collection							
Withdraw and dropout							
Analyses							
Intervention integrity							
Controls were better educated had greater number of full-term pregnancies, use oral contraceptive more, higher frequency of tubal ligation							

West, 1966 [UI# 5939299]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Cases: 53 Controls: 50 Menopause: mix Race: ND Enrolled/Evaluate: Cases: 97/76 Controls: 91/76 Location: USA Sites: Multi Funding: ND	Cases: Pseudo-mucinous cystadenocarcinoma 23 Serous cystadenocarcinoma 16 Cystadenocarcinoma and adenocarcinoma 43 Carcinoma 6 Granulosa cell 5 Teratoma 1 Dysgerminoma 1 Meso-meta-nephroma 1 Unclassified 1 Controls: (25 diagnoses) Endometriosis 19 Simple cyst 10 Pseudo-mucinous cystadenoma 9 Dermoid cyst 8	Case-control Data collection by interview	Cases: malignancy of ovary, 50 mile radius of Boston, exclusion includes >75 years and co-existent malignancy of another organ that was not metastatic from ovary and recurrent cases Controls: female hospital patients matched for age $\pm$ 5 years, residence and date of surgery, next operated case of benign ovarian neoplasms from same hospital as malignant case	Breast feeding status determined by interview	Breast feeding status determined by interview

West, 1966 [UI# 5939299]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Malignant ovarian cancer confirmed by pathology	T test with p level of 0.01 for significance	76 matched pairs Lactation (months) Cases Controls 6.6 5.8 p > 0.3	A: strong, B: moderate, C: weak	A	B	C
			Selection			X
			Study design		X	
			Confounder			X
			Blinding			X
			Data collection		X	
			Withdraw and dropout		X	
			Analyses			X
			Intervention integrity			
			Cases and controls classified as private and service with more control taken from service care implying differ SES. Only descriptive statistics given for education. Not adjusting for potential confounders			

Whittemore, 1992 [UI# 1476141]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Cases: ND Controls: ND Menopause: mix Race: White Enrolled/Evaluate: Cases: 2197/1071 Controls: 8893/5705 Location: USA Sites: 12 studies, of which 7 with breastfeeding data Funding: Government	ND	Case-control Data from 12 studies conducted from 1956-1986. 6 hospital and 6 population (random-digit dialing) controls. Breastfeeding data from 7 studies; 2 hospital- and 5 population-based studies	Studies with individual patient data collected from personal interviews through structured questionnaires, data coded and stored electronically, variable definitions documented Cases: newly diagnosed invasive epithelial tumors at US hospital Controls: reside in US, during time of case ascertainment, no gynecological condition if hospital controls	Breastfeeding status determined by interview	Breastfeeding status determined by interview

Whittemore, 1992 [UI# 1476141]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Invasive epithelial tumors	Conditional logistic regression with odds ratio adjusted for age, study, parity, oral contraceptive use	Breastfeeding duration (month)*	Cases n (%)	Controls n (%)	RR _{adj} (95% CI)	A: strong, B: moderate, C: weak	A	B	C
		Hospital studies				Selection		X	
		0	117 (58)	612 (57)	1.0	Study design		X	
		1-5	49 (24)	237 (22)	0.95 (0.62-1.5)	Confounder			X
		6-11	13 (6)	110 (10)	0.40† (0.20-0.78)	Blinding		X	
		12-23	16 (8)	78 (7)	0.70 (0.36-1.4)	Data collection			X
		≥24	6 (3)	44 (4)	0.59 (0.22-1.6)	Withdraw and dropout		X	
		Any breastfeeding	84 (42)	469 (43)	0.73 (0.51-1.0)	Analyses		X	
		Trend per month			0.99 p=0.18	Intervention integrity			
		Population studies							
		0	478 (55)	2152 (47)	1.0				
		1-5	208 (24)	1188 (26)	0.87 (0.72-1.1)				
		6-11	88 (10)	567 (12)	0.74 (0.57-0.96)				
		12-23	61 (7)	449 (10)	0.69 (0.51-0.94)				
		≥24	35 (4)	268 (6)	0.74 (0.49-1.1)				
		Any breastfeeding	392 (45)	2472 (53)	0.81† (0.68-0.95)				
		Trend per month			0.99†				
		* Parous † p<0.01 Risk reduction per month of breastfeeding within 6 months delivery exceeds that for subsequent breastfeeding Hospital 2.5% vs 1.4% Population 1.2% vs 0.9%							

Hospital-based studies overlap with Hartge 1989

Population-based studies overlap with Cramer 1983

Wynder, 1969 [UI# 5764976]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Cases: 52 Controls: ND Menopause: mix Race: Cases: 91% White 9% Black Controls: 94% White 6% Black Enrolled/Evaluate: Cases: 158 Controls: 300 Location: USA Sites: Multi Funding: Government	Serous cystadenocarcinoma 28 Pseudo-mucinous cystadenocarcinoma 10 Endometrioid carcinoma 11 Solid adenocarcinoma 91 Miscellaneous Granulose cell tumor 4 Teratoma 1 Carcinoid 1 Others 2 Unspecified 10	Case-control Hospital-based population Data collection by interview	Epithelial ovarian cancer confirmed by histology, functionary or dysontogenetic tumors excluded Eight additional cases of diagnosis other than adenocarcinoma and endometrioid cancers Controls matched for age	Breastfeeding status determined by interview	Breastfeeding status determined by interview

Wynder, 1969 [UI# 5764976]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments				
Epithelial ovarian cancer confirmed by histology	ND	No difference between groups for age of first nursing or never nursing				A: strong, B: moderate, C: weak	A	B	C	
			158 Cases		300 Controls		Selection			X
			Premenopausal (55)	Postmenopausal (95)	Premenopausal (95)	Postmenopausal (205)	Study design			X
		BF ≥12 months	10%	24%	7%	21%	Confounder			X
						Blinding		X		
						Data collection			X	
						Withdraw and dropout		X		
						Analyses			x	
						Intervention integrity				

Yen, 2003 [UI# 12713998]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Median age (range): Cases: 47 (20-75) Controls: 44 (20-75) Menopause: mix Race: ND Enrolled/Evaluate: Cases: 86 Controls: 369 Location: Taiwan Sites: Multi Funding: Government	Cases: Serous (27%) Mucinous (27%) Endometrioid (21%) Clear cell (15%) Unspecified adenocarcinoma (2%) Controls: Trauma (28%) Nontraumatic orthopedic conditions (30%) Surgery (19%) Miscellaneous (23%) medical, eye, nose, throat, dental	Case-control Data collection by in-hospital interview	Newly diagnosed primary invasive epithelial ovarian cancer confirmed by histology, Taiwan resident for > 20 years Controls: matched for age, same hospital admission at same time for nonmalignant, nongynecologic condition, nonhormonal or nondigestive tract diseases, no long term modification of diet Exclusion: major gynecologic operation including hysterectomy with or without oophorectomy or oophorectomy only, severe illness	Breast feeding status determined by interview	Breast feeding status determined by interview

Yen, 2003 [UI# 12713998]

Outcome Definition	Statistical analyses and confounders adjusted	Results				Bias/limitations Comments			
Primary invasive epithelial ovarian cancer confirmed by histology	Conditional multivariate logistic regression model adjusted for number of live birth	Breastfeeding duration (years)	Cases n (%)	Controls n (%)	OR _{adj} (95% CI)	A: strong, B: moderate, C: weak	A	B	C
		0	41 (48)	176 (48)	1.0	Selection		X	
		≤1	8 (9)	42 (11)	0.82 (0.35-1.9)	Study design		X	
						Confounder		X	
		>1	37 (43)	151 (41)	0.55 (0.29-1.01)	Blinding			X
		Breastfeeding for more than one year is trend for protection				Data collection		X	
						Withdraw and dropout		X	
						Analyses		X	
						Intervention integrity			

Appendix C. Evidence Tables

Part III. Maternal Outcomes

Postpartum Depression

Chaudron 2001
Cooper 1993
Hannah 1992
Henderson 2003
Murray 2003
Seimyr 2004
Warner 1996

Chaudron 2001, [UI# 1446151]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 29 yr (SD 4.4) Race: 69% white Enrolled/Evaluate: 465 Location: US Sites: Multi Funding:		Prospective cohort; the study examined predictors of new-onset postpartum depression during postpartum months 1-4; data from the Wisconsin Maternity Leave and Health Project (WMLHP); subjects participated in a 90-minute home interviews during the second trimester, and at 1 and 4 months after delivery; both the DIS and CES-D were administered at each interview; DIS questions about fatigue and sleep were modified to account for disturbances caused by night-time infant care	Between weeks 12 and 25 of a non-high-risk pregnancy; >18 yr; non-handicapped; living within Milwaukee or Dane counties; living with the baby's father; one of the couple was employed; had a telephone; spoke English well enough to understand interviewer; able to complete questionnaire; not depressed at 1 month postpartum; Excluded subjects who did not complete follow up at 4 months		

Chaudron 2001, [UI# 1446151]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Postpartum depression defined by: 1. Diagnosis of major depression on NIMH Diagnostic Interview Schedule (DIS) per DSM-III-R criteria 2. ≥ 16 on the Center for Epidemiologic Studies Depression Scale (CES-D); and /or 3. receiving antidepressants	Age, depression during pregnancy, postpartum thoughts of death and dying, difficulty falling asleep (see full discussion in paper)	327/465 breastfed; 27/465 became depressed between 1 and 4 months; Women who breastfed their infants were not significantly different in their development of depression from women who bottlefed their infants. Of those women who were breastfeeding at 1 month postpartum (n=327), women who worried about breastfeeding were significantly more likely to become depressed than those who did not worry (relative risk 3.0, CI 1.041-9.216, P=0.04)	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder		x	
			Blinding			x
			Data collection		x	
			Withdraw and dropout			
			Analyses		x	
			Intervention integrity			

Cooper, 1993 [UI# 8463993]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Enrolled/Evaluate: 243 women at Oxford; 113 at Cambridge Location: UK Sites: Multi Funding:		2 separate prospective cohort studies (Oxford and Cambridge)	See ref 12 in paper for criteria for Oxford; For Cambridge: primiparous, 20-40 yr, had a partner, 37-42 wk gestation, infant $\geq$ 2.5 Kg, and no gross congenital abnormality	Of those who attempted to initiate breastfeeding, those who terminated before 8 wk were compared with those who did not	

Cooper, 1993 [UI# 8463993]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments																																				
At Oxford, Present State Examination (PSE) was used to determine psychiatric symptoms. At Cambridge, prospective subjects were screened at 6 wk postpartum using the Edinburgh Postnatal Depression Scale (EPDS). All high scorers (≥ 13), 78% of the intermediate scorers (10-12), and a random sample of low scorers (≤ 9) received full psychiatric assessment between 2 and 3 months postpartum. 58 who met Research Diagnostic Criteria (RDC) criteria for major and minor depression, and 55 who were psychiatrically well were interviewed in full. At Cambridge, Standardized Psychiatric Interview (SPI) was used.	Social class, age, and education	At Oxford, 14/25 subjects (56%) with psychiatric symptoms versus 40/175 subjects (23%) without psychiatric symptoms had given up breastfeeding by 8 wk postpartum ($P < 0.01$). In eight, depression preceded cessation of breastfeeding, in two depression was the subsequent event, and in four, the two events arose contemporaneously. Low social class and younger age were also determinants of early cessation of breastfeeding. Antenatal psychiatric status did not predict breastfeeding outcome. At Cambridge, 30/54 subjects (56%) with an episode of depression versus 10/48 subjects (21%) without depression postpartum had given up breastfeeding by 8 wk ($P < 0.001$). Lower educational attainment was also associated with early cessation of breastfeeding. In all cases, the onset of depression preceded the cessation of breastfeeding.	<table border="1"> <tr> <td>A: strong, B: moderate, C: weak</td><td>A</td><td>B</td><td>C</td></tr> <tr> <td>Selection</td><td></td><td>x</td><td></td></tr> <tr> <td>Study design</td><td></td><td></td><td>x</td></tr> <tr> <td>Confounder</td><td></td><td>x</td><td></td></tr> <tr> <td>Blinding</td><td></td><td></td><td>x</td></tr> <tr> <td>Data collection</td><td></td><td></td><td>x</td></tr> <tr> <td>Withdraw and dropout</td><td></td><td></td><td>x</td></tr> <tr> <td>Analyses</td><td></td><td>x</td><td></td></tr> <tr> <td>Intervention integrity</td><td></td><td></td><td></td></tr> </table>	A: strong, B: moderate, C: weak	A	B	C	Selection		x		Study design			x	Confounder		x		Blinding			x	Data collection			x	Withdraw and dropout			x	Analyses		x		Intervention integrity			
A: strong, B: moderate, C: weak	A	B	C																																				
Selection		x																																					
Study design			x																																				
Confounder		x																																					
Blinding			x																																				
Data collection			x																																				
Withdraw and dropout			x																																				
Analyses		x																																					
Intervention integrity																																							

Hannah 1992, [UI# 1617360]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Evaluated: 231 Location: UK Sites: Multi Funding:	ND	Prospective cohort, questionnaires (general questionnaire and Edinburgh Postnatal Depression Scale (EPDS)) on 5 th day postpartum and repeat EPDS at 6 weeks postpartum	Women who had delivered a live baby	ND	

Hannah 1992, [UI# 1617360]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Postpartum depression defined by EPDS score ≥ 13 at 6 weeks.	Chi-square	At 6 weeks, 8/26 (31%) women with EPDS score ≥ 13 vs 25/200 (12%) women with EPDS < 13 only did bottle feeding ($P < 0.04$). [Or 18/26 (69%) women with EPDS score ≥ 13 vs 175/200 (88%) women with EPDS < 13 had ever breastfed.]	A: strong, B: moderate, C: weak	A	B	C
			Selection			x
			Study design			x
			Confounder			x
			Blinding			x
			Data collection			x
			Withdraw and dropout			x
			Analyses			x
			Intervention integrity			

Henderson 2003, [UI# 12911800]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Race: Enrolled/Evaluate: 1745 enrolled Location: Australia Sites: Multi Funding:		Prospective cohort, population from a RCT on the effectiveness of maternal debriefing on the incidence of postnatal psychological morbidity; self-reported questionnaires were completed at 2, 6, and 12 months after birth; the questionnaire included a measure of current breastfeeding status; The Edinburgh Postnatal Depression Scale (EPDS) was used to screen for symptoms of depression at each of the follow-up intervals. All subjects who scored > 12 on the scale, subjects being treated for a psychological disorder, subjects taking antidepressants, and a stratified sample of women with low EPDS scores were offered a diagnostic interview. The interview enabled a diagnosis of depression to be made based on DSM IV.	Before 72 hr postpartum, given birth to health infants >35 wk gestation, English speaking, >18 yrs, not under psychological care at the time of recruitment; cases were excluded if no information about breastfeeding status was given at any interval (n=56)	Primary endpoint was duration of breastfeeding	

Henderson 2003, [UI# 12911800]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Diagnosis of depression based on DSM IV (data obtained from diagnostic interview)	Adjusted for demographic, perinatal, and postnatal factors	Breastfeeding was initiated by 1670/1745 women (96%); 57% still breastfeeding at 6 months (30% fully); 22% still breastfeeding to some extent at 12 months. 314/1745 (18%) developed depression in the 12 months after birth, 63% showing the first symptom by 2 months. After adjustment for confounding factors, early cessation of breastfeeding was found to be associated with postnatal depression (adjusted hazard ratio 1.25, 95% CI 1.03 to 1.52). Onset of postnatal depression occurred before cessation of breastfeeding in most cases.	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder		x	
			Blinding			x
			Data collection		x	
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity			

Murray, 2003 [UI# 12848402]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 25 yr in self-exclusion group; 28 yr in take-up group Menopause: pre/post/mix Race: 75 in self-exclusion group; 81 in take-up group Location: UK Sites: Single/Multi Funding:	History of depression was common, as were health problems, and depressed or anxious mood in pregnancy	Cohort study; 403 women from a RCT of a preventive intervention for postpartum depression were identified as being high risk for depression from antenatal screening; 2 groups were formed; one group from the control arm (routine care) of the RCT was willing to receive additional home-based Health Visiting in the last 5 weeks of pregnancy and the first 8 weeks postpartum; the other group refused the intervention	Vulnerable to postpartum depression per screening results		

Murray, 2003 [UI# 12848402]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Predictive Index for Postpartum Depression		Comparing the self-exclusion group (n=75) with the take-up group (n=81), the infants in the self-exclusion group were less likely to be breastfed, both immediately after birth (63% vs 83%, P=0.007) and at 10 days (31% vs 54%, P=0.01)	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design			x
			Confounder			x
			Blinding			x
			Data collection			x
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity			

Seimyr 2004, [UI# 15376402]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 29 yr (18-43) Race: Enrolled/Evaluate: 352 women Location: Sweden Sites: Multi Funding:		Prospective cohort, Edinburgh Postnatal Depression Scale (EPDS) result at 2 months before childbirth (I), 2 months (II) and 12 months (III) after childbirth	All Swedish speaking pregnant women at six antenatal clinics were invited to participate, with their partners or alone		

Seimyr 2004, [UI# 15376402]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Cut-off point of 9/10 on EPDS scale is used to set the threshold for vulnerability to depression	Did not report actual adjustment, but reported that there was no difference with regards to age, education, parity, and length of marital relationship between low and high-scoring women	63 women (21%) scored high on EPDS at the end of pregnancy; 54 (17%) at 2 months; 28 (12%) at 12 months. ~50% of the high EPDS (I) scorers also scored high on EPDS (II), and 39% of those who were high EPDS (II) scorers continued to score high on EPDS (III). Fewer high EPDS (I) scorers breastfed compared to the low scorers (82% vs 94%, P<0.02) Fewer high EPDS (II) scorers breastfed compared to the low scorers (85% vs 93%, P<0.08) High EPDS scorers experienced breastfeeding more negatively than the women scoring low on EPDS (I) (40% vs 20%, P<0.03) and on EPDS (II) (51% vs 16%, P<0.0001). High EPDS (II) scorers breastfed for a shorter time compared to the low-scoring women (4.6 months, SD 2.4 vs 5.3 months, SD 1.9, P<0.04)	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design			x
			Confounder			x
			Blinding			x
			Data collection		x	
			Withdraw and dropout		x	
			Analyses		x	
			Intervention integrity			

Warner, 1996 [UI# 8733800]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 28 yr (15-46 yr) Race: Eligible/Enrolled: 2978/2375 Location: UK Sites: Multi Funding:	Prospective cohort	Recruited subjects over a 20-month period from two postnatal wards prior to discharge, subjects agreed to a home visit 6-8 wk after delivery; Edinburgh Postnatal Depression Scale (EPDS) was completed at the interview	Living outside the district and insufficient English to understand the questionnaire		

Warner, 1996 [UI# 8733800]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
EPDS score > 12	13 socio-demographic and obstetric variables were entered individually into a logistic regression analysis with high and low EPDS as the outcome variable using a threshold of 12/13; those variables showing a significant association with high EPDS scores were entered into a stepwise logistic regression analysis	280/2375 scored >12 on EPDS; Unplanned pregnancy, not breastfeeding at 6 wk (OR 1.52, 95%CI 1.12-2.06), unemployment in mother or head of household were associated with an EPDS score >12 in a stepwise logistic analysis	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design			x
			Confounder	x		
			Blinding			x
			Data collection		x	
			Withdraw and dropout		x	
			Analyses			x
			Intervention integrity			

Appendix C. Evidence Tables

Part III. Maternal Outcomes

Maternal Weight

Brewer 1989
Janney 1997
Linne 2003
Ohlin 1990
Olson 2003
Sichieri 2003
Walker 2004

Brewer 1989 [2916446]

Study characteristics	Population	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Pre-pregnancy BW (range): ND Pre-pregnancy BMI (range): ND Race: white Enrolled/Evaluate: 70/56 Location: US Sites: Single Funding:	Well-educated, middle to upper middle income range	Prospective cohort; to study postpartum changes in body weight in lactating vs. nonlactating women up to 6 mo postpartum; mothers were visited in the hospital within 1 to 2 d of delivery and at home at 3 and 6 mo postpartum, weights were objectively measured; detailed information on infant diet was collected by monthly questionnaires completed by the mothers. Mother's calorie intake measured. Three-day maternal food records were also completed at 3 and 6 months postpartum. Energy intake (kcal) was calculated by computer from food records with the Nutritional Analysis System (Louisiana State University, Baton Rouge, LA).	Good health, $\geq$ 18 yr, mothers of full term infants		

Brewer 1989 [2916446]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Exclusive breastfeeding; exclusive formula feeding; mixed breast and formula feeding	ANCOVA; covariables included maternal age, parity, prepregnancy weight, initial postpartum measurement for each variable, and others	“If weight change during the first 3 mo is corrected for an initial fluid loss of 2 kg, weight loss averaged 1.6, 2.1, and 1.5 kg/mo for exclusive breastfeed, exclusive formula feed, and mixed feed, respectively. Exclusive breastfeed resulted in an additional 0.43 kg/mo loss over the second 3 mo with mixed feed averaging a 0.27 kg loss and nonlactating women experiencing essentially no change.” No significant differences in total weight loss were observed between the lactating and non-lactating groups. Exclusive breastfeeding resulted in a significant decrease in weight between 3 and 6 months (P < 0.05).	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder		x	
			Blinding			x
			Data collection		x	
			Withdraw and dropout			x
			Analyses		x	
			Intervention integrity			
Overall: B						

Janney, 1997 [9356528]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 29.3 (20-40) Pre-pregnancy BW (range): 59.7 (43.1-93.0) kg Pre-pregnancy BMI (range): 22.2 (16.9-33.7) kg/m ² Race: 96% White Enrolled/Evaluate: 110/71 Location: US Sites: Single Funding: ND		Prospective, longitudinal study (0.5, 2, 4, 6, 12, and 18 mo post-partum)	Pregnant nulliparous and primiparous women aged 20-40 y in their 3 rd trimester were recruited from birthing education classes and obstetric practices. To be eligible for participation, the women must have declared either that they had no intention of breastfeeding or that they intended to breast-feed for ≥ 6 months. Exclusion criteria were a history of endocrine, renal, liver, or chronic respiratory illness; complications of pregnancy, including HTN and GDM; fetal complications or complications at delivery; delivery of an infant small for gestational age; and delivery of twins.	Fully breast-feeding was defined as providing at least 2/3 of the needed energy intake per kilogram of the infant's weight in breast milk. See article for how this estimation was made.	Partly breast-feeding or bottle-fed

Janney, 1997 [9356528]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments
Postpartum weight retention over time	Longitudinal regression analysis: proc mixed and glm procedures	Most women (66.7%) fully breastfed for 6 mo. Duration of lactation practice was a significant predictor of postpartum weight retention over time - ($P < 0.001$) - not only in the longitudinal regression model containing only lactation practice and months since parturition but also in the multiple-variable longitudinal regression model. Regardless of lactation practice, women consistently lost weight from 0.5 to 18 mo after parturition. The weight-retention curve appeared to be curvilinear, with slower rates of weight loss occurring after 12 mo postpartum. Overall, less weight was retained by lactating women than by nonlactating women. Even women who breast-fed for < 4 mo retained less weight than women who bottle-fed their infants. Additionally, women who switched to partly breast-feeding had weight retention rates intermediate between those of women fully breast-feeding and those bottle-feeding. Once lactation was discontinued, slower rates of weight loss were observed. Breast-feeding women achieved their prepregnancy weights about 6 month earlier than women who only bottle-fed their infants (see Figure 5 in original paper).	

Janney, 1997 [9356528]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
		The following data were <u>estimated from Figure 5 in the original paper</u>: 1) Time returned to prepregnancy weights: Bottle-feeding only: ~18 months postpartum Partly breast-feeding at 2 mo, and bottle-feeding or infant weaned at 4, 6, 12, and 18 months: ~12 months postpartum Fully breast-feeding for 6 month and bottle-feeding or infant weaned at 12 and 18 month: ~11 months postpartum Fully breastfeeding for 6 month, partly breast-feeding for 12 month, and bottle-feeding or infant weaned at 18 month: ~ 10 months postpartum 2) Weight retention at 6 months postpartum: Fully breastfeeding for 6 months (n=57): ~ +3 kg Partly breastfeeding at 2 month and bottle-feeding at 4 and 6 months (n=10): ~ +3.5 kg Fully bottle-feeding for 6 months (n=12): ~ +5 kg	A: strong, B: moderate, C: weak Selection Study design Confounder Blinding Data collection Withdraw and dropout Analyses Intervention integrity Overall: B	A	B	C
					x	
				x		
						x
				x		
						x
					x	

Linne, 2003 [14634683]

Study characteristics	Population	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 45 yr (at 15 yr follow up) Pre-pregnancy BW (range): 59.8 kg Pre-pregnancy BMI (range): 21.5 kg/m ² Race: Enrolled/Evaluate: 1423/563 (at 15 yr) Location: Stockholm Sites: Multi Funding:	Cross section of the metropolitan population from both inner city of Stockholm and its suburbs; a range of socioeconomic groups	Prospective cohort; longitudinal study of women's weight gain during pregnancy, at 1 yr postpartum and at 15 yr follow up; feeding data were collected retrospectively by questionnaires at 2.5, 6, 12 mo and at 15 yr; for women who still lived in the Stockholm area, weights were measured in the University hospital (n=363); for women who lived far away, weights were self-reported (n=200)	2 outliers were excluded: BMI of 47; first child at 49 yr	Measured by lactation score: every month of full lactation was given 4 points, every month of mixed lactation was given 2 points; total sum (0-48) indicated the individual lactation amount	

Linne, 2003 [14634683]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Overweight (BMI > 25), normal weight (BMI 20-25);	ANOVA was used to analyze weight change at four time points; women who were overweight at prepregnancy and at follow up and those who lost weight and regained a BMI in the normal range at 15 yr were excluded from the analysis	Those women who became overweight had lower lactation scores than women who remained normal weight at 15 years follow up (21.7 ± 11.0 vs. 24.0 ± 9.4 , $t=2.25$, $df = 488$, $P < 0.05$). Responders were slightly older, had higher educational attainment, and higher income than non-responders. Nonresponse was more common in non-Nordic citizens.	A: strong, B: moderate, C: weak Selection Study design Confounder Blinding Data collection Withdraw and dropout Analyses Intervention integrity Overall: C			

Ohlin, 1990 [2341224]; Ohlin, 1996 [8732961]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 29.5 (17-49) Pre-pregnancy BW (range): 59.6 (39-129) kg Pre-pregnancy BMI (range): 21.5 (15.4-43.0) kg/m ² Race: ND Enrolled/Evaluate: 2295 (at 2 mo) /1423 (at 1 yr) Location: Sweden Sites: Multi Funding: Government		Women were studied retrospectively during pregnancy at maternity clinics, and monitored prospectively up to 1 year after delivery.	All mothers who, during one year, came to the maternity clinic for the last routine control, i.e. 6-15 weeks after the delivery, were invited by their midwives to take part in the study. Women who were going to move within a short time and those with obvious language and communication problems were not invited. Exclusion: twin births, use of insulin during pregnancy, GI problems with severe energy losses (heavy vomiting or diarrhea) and pre-pregnant body weights were not available.	Lactation score: a scoring system was constructed in order to express duration and intensity of breast feeding. Every month with full lactation was given 4 points, and every month with mixed feeding was given 2 points. The total sum (0-48) indicated the individual lactation amount.	

Ohlin, 1990 [2341224]; Ohlin, 1996 [8732961]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
The body weight 1 year after delivery (Δ weight)	Ranged influence of different factors on the Δ weight was analyzed by a multiple stepwise regression analysis	30% lost weight, 56% gain 0-5 kg, 13% gained 5-10 kg and 1.5% gained ≥ 10 kg 1 year after delivery. Frequency of overweight women (BMI ≥ 23.9) increased from 13% before pregnancy to 21% 1 year post-partum (p<0.001) Ranged influence of different factors on the Δ weight: The coefficients between the Δ weight and lactation score, age, pre-pregnancy weight and parity were very low (multiple r=0.38, p<0.001) and significant correlations do not necessarily indicate causality. Weight gain during pregnancy had the strongest influence, and explained 12.7% of the variation of the Δ weight (p<0.001). Lactation and age increased the total proportion of explained variance by about 1% each. Simple correlation coefficient between the Δ weight and lactation score was -0.09 (p<0.01)	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder		x	
			Blinding			
			Data collection			x
			Withdraw and dropout			x
			Analyses		x	
			Intervention integrity			

Olson, 2003 [12532163]

Study characteristics	Population	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): Pre-pregnancy BW (range): kg Pre-pregnancy BMI (range): kg/m ² Race: white (96%) Enrolled/Sample/Evaluate: 622/597/540 Location: US Sites: Single Funding:	Rural, socio-economically diverse	Prospective cohort; to study the relative importance of gestational weight gain, postpartum exercise, food intake and breastfeeding to weight change from early pregnancy to 1 yr postpartum; weight was measured at antenatal and 1 yr postpartum visit (varied between 9 and 19 mo); 32/540 self-reported weight at 1 yr; breastfeeding data collected at 6 wk obstetric visit; 6 mo and 1 yr questionnaires. Food frequency questionnaires were completed at 1 st or 2 nd trimester of pregnancy, 6 months postpartum, and 1 yr postpartum.	≥ 18 yr; healthy, gave birth to full-term singletons	Breastfeeding after 6 mo was considered to be non-exclusive; a breastfeeding score similar to Ohlin and Rossner's was constructed: 1 point for each wk of exclusive breastfeeding and 0.5 point for each wk of mixed feeding	

Olson, 2003 [12532163]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments																																				
Weight change from early pregnancy to 1 yr postpartum; major weight gain was defined as ≥ 4.55 kg at 1 yr postpartum	Univariate and multiple linear regression; weights taken between 9 and 19 mo were eligible for inclusion, actual number of mo postpartum at the measurement of the "1 yr postpartum body weight" was entered into all of the regression models; variables in the initial model included age, education, smoking status, income, marital status, prepregnancy BMI category and parity; 3 subjects with poor fit were removed from the analysis; a model reduction method was applied at the 10% significance level to produce an inclusive, reduced model Using food frequency data, the difference in the daily energy intake from 6 months to 1 yr postpartum was calculated for each woman. The Wilcoxon rank sum test was used to assess change in total daily energy intake across the five categories of self-reported behavior change in amount of food intake.	Only the 597 women who delivered full term infants were included in the sample; 540 had 1 yr weight recorded; Women who were breastfeeding at 1 yr retained less weight compared with the women who weren't (P = 0.04). Breastfeeding at all other time points and the breastfeeding score were not significantly related to postpartum weight retention.	<table border="1"> <tr> <td>A: strong, B: moderate, C: weak</td><td>A</td><td>B</td><td>C</td></tr> <tr> <td>Selection</td><td></td><td>x</td><td></td></tr> <tr> <td>Study design</td><td></td><td>x</td><td></td></tr> <tr> <td>Confounder</td><td></td><td>x</td><td></td></tr> <tr> <td>Blinding</td><td></td><td></td><td>x</td></tr> <tr> <td>Data collection</td><td></td><td>x</td><td></td></tr> <tr> <td>Withdraw and dropout</td><td></td><td>x</td><td></td></tr> <tr> <td>Analyses</td><td></td><td>x</td><td></td></tr> <tr> <td>Intervention integrity</td><td></td><td></td><td></td></tr> </table> Overall: B	A: strong, B: moderate, C: weak	A	B	C	Selection		x		Study design		x		Confounder		x		Blinding			x	Data collection		x		Withdraw and dropout		x		Analyses		x		Intervention integrity			
A: strong, B: moderate, C: weak	A	B	C																																				
Selection		x																																					
Study design		x																																					
Confounder		x																																					
Blinding			x																																				
Data collection		x																																					
Withdraw and dropout		x																																					
Analyses		x																																					
Intervention integrity																																							

Sichieri, 2003 [12821967]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 24-40 yr Pre-pregnancy BW (range): Stratified by BMI at baseline Pre-pregnancy BMI (range): % BMI $\leq 25\text{kg/m}^2$ in 1989: 81.5% BW mean, ALL: 58.0 kgs % BMI $> 25\text{kg/m}^2$ in 1989: 18.5% BW mean, ALL: 80.4 kgs Race: ND Enrolled/Evaluate: 1,538/1,538 nulliparous at baseline; 2,810/2,810 primiparous at baseline Location: USA Sites: Single/Multi: mailed questionnaire Funding: Supported by FundaCBo CAPESCoordenaq20 de Aperfeicoamento de Pessoal de Nivel Superior and Boston Obesity Nutrition Research Center (DK 46200)	None documented	The cohort of the Nurse's Health Study II, with analysis restricted to women who were aged 24 to 40 y at baseline (1989), who had a history of no more than one past full-term pregnancy at baseline, gave birth to one child between 1990 and 1991, but had no other pregnancies during the follow-up. SUBJECTS: 1,538 of the 33,082 nulliparous women and 2,810 of the 20,261 primiparous, in 1989. The NHS II Cohort has been followed up every 2 yr to ascertain incident diseases and exposures including parity and body weight. Individual breastfeeding data were collected retrospectively in the 1997 follow-up. Women were asked to recall their lifetime breastfeeding history.	Excluded from analysis: those who reported use of insulin, hypoglycemic drugs, or thyroid disease at baseline or followup were excluded from analysis	For each birth, women were asked if they had breastfed for at least a month and if so they were asked to report the month in which they introduced the formula. Introduction of daily formula/milk was assumed to represent the end of exclusive breastfeeding period. Duration of exclusive breastfeeding was categorized into 0, 1-3, 4-7, 8-11, and 12 months or more. <u>Exclusive breastfeeding:</u> Women were asked to recall their lifetime breastfeeding history. They were asked to answer 2 questions: 1) "Did you breastfeed at least one month?" 2) "Which month did you introduce formula?" If the answer to the 1st questions was "no, not at all", the duration was considered to be zero. Data from these 2 questions were combined with parity data from 1991 and 1993 questionnaires.	

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments																																				
To assess whether women who breastfed were different in terms of their weight changes than women who did not, weight change from prepregnancy (in 1989) to postpregnancy (in 1993) was estimated according to exclusive breastfeeding duration using linear regression models.	The longitudinal models were based on mixed effects analysis. These models permitted the authors to test differences in prepregnancy weight at baseline (group effect), weight gain with pregnancy (time-dependent effect of period I), and weight change after pregnancy (time-dependent effect of period II) by breastfeeding status. If breastfeeding groups differed at baseline but had the same change in weight during follow-up, a group effect but not a time-dependent effect would be seen.	After adjusting for age, physical activity, and BMI, lactation was associated with a weight gain from 1989 to 1993 of approximately 1 kg (statistically significant only for women nulliparous in 1989 with a BMI <25 kg/m² (P=0.02) and for those women primiparous in 1989, with a BMI ≥25 kg/m² (P=0.04)) comparing women who breastfed with women who did not. Duration of lactation was unrelated to the magnitude of weight change (P>0.40 for all comparisons). Maternal weight change associated with breastfeeding is minimal among normal weight women, and for overweight women the weight gained in pregnancy is not reduced by breastfeeding.	<table border="1"> <thead> <tr> <th>A: strong, B: moderate, C: weak</th><th>A</th><th>B</th><th>C</th></tr> </thead> <tbody> <tr> <td>Selection</td><td>x</td><td></td><td></td></tr> <tr> <td>Study design</td><td></td><td>x</td><td></td></tr> <tr> <td>Confounder</td><td>x</td><td></td><td></td></tr> <tr> <td>Blinding</td><td></td><td></td><td>x</td></tr> <tr> <td>Data collection</td><td></td><td>x</td><td></td></tr> <tr> <td>Withdraw and dropout</td><td></td><td>x</td><td></td></tr> <tr> <td>Analyses</td><td>x</td><td></td><td></td></tr> <tr> <td>Intervention integrity</td><td></td><td></td><td></td></tr> </tbody> </table> Overall: B	A: strong, B: moderate, C: weak	A	B	C	Selection	x			Study design		x		Confounder	x			Blinding			x	Data collection		x		Withdraw and dropout		x		Analyses	x			Intervention integrity			
A: strong, B: moderate, C: weak	A	B	C																																				
Selection	x																																						
Study design		x																																					
Confounder	x																																						
Blinding			x																																				
Data collection		x																																					
Withdraw and dropout		x																																					
Analyses	x																																						
Intervention integrity																																							

Walker, 2004 [15778139]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): ND (>18 yr) Pre-pregnancy BW (range): ND Pre-pregnancy BMI (range): ND Race: 26% African American, 44% Hispanic, 30% White Enrolled/Evaluate: ND/382 Location: US Sites: Single Funding: Government	healthy	Prospective, longitudinal cohort (post-delivery, and 6 wk, and 3, 6, and 12 months postpartum). Energy intakes were measured at above times by average of one 24-hour recall and two food records at each observation. Fat intakes measured by Food Habits Questionnaire (Kristal, 1990).	Healthy women, singleton pregnancies with prenatal care funded through Medicaid; had a parity not exceeding III; were conversant in English; and were age 18 or higher.	Full (exclusive) or partial breastfeeding, or bottle feeding, measured at post-delivery, 6 wk, and 3, 6, and 12 months postpartum	

Walker, 2004 [15778139]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Postpartum BMI	General linear mixed models, also known as hierarchical linear models, were used. Forward additive method (or forward selection) was used to select variables for inclusion in the model.	Infant feeding method was not associated with postpartum BMI (p=0.140) when ethnicity, time, pre-pregnant BMI and a bunch of other variables (such as parity, gestational weight gain, Cesarean section, etc) were included in the model.	A: strong, B: moderate, C: weak	A	B	C
			Selection		x	
			Study design		x	
			Confounder		x	
			Blinding			
			Data collection		x	
			Withdraw and dropout			x
			Analyses		x	
			Intervention integrity			
			Overall: B			

Appendix C. Evidence Tables

Part III. Maternal Outcomes

Type 2 DM

Stuebe 2005

Stuebe, 2005 [16304074]

Study characteristics	Subjects' health conditions	Study design and follow-up duration	Eligibility criteria	Breastfeeding Exposures or Interventions	Control Exposures or Interventions
Mean age (range): 52 at follow-up Menopause: mix Race: ND Enrolled/Evaluate: see results Location: US Sites: Multi Funding: Government	ND	Prospective, longitudinal study	The Nurses' Health Studies consist of 2 large cohorts enrolled in prospective, longitudinal studies of women's health The Nurses' Health Study (NHS) was initiated in 1976 and enrolled 121,700 women from 11 states. Participants were between 30 and 55 years of age at baseline, and each woman completed a detailed baseline questionnaire regarding diseases and health related topics. The second cohort, the Nurses' Health Study II (NHS II), began in 1989, enrolling 116,671 women from 14 states. Participants were between 25 and 42 years of age and completed a similar baseline questionnaire as well as biennial follow-up questionnaires.	Women in the Nurses' Health Studies (NHS) reported parity at baseline in 1976. Lactation history was assessed once, in 1986, when women were asked to report total lifetime duration lactation. In NHS II, women reported the number of pregnancies lasting more than 5 months at baseline and on each biennial questionnaire. Lactation history was assessed 3 times. In 1993, 1997, and 2003. Information on parity was used to derive retrospectively each participant's total cumulative lactation at each 2-year interval. Total duration of lactation was calculated based on the number of months after birth that the participant reported stopping breastfeeding altogether. Using the 1997 and 2003 NHS II data, duration of exclusive lactation could be calculated based on the reported timing of introduction of formula or solid food.	

Stuebe, 2005 [16304074]

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments			
Ascertainment of Type 2 Diabetes Mellitus A case of diabetes was confirmed if a woman reported 1 of the following: (1) 1 or more classic symptoms plus either a fasting glucose of ≥ 140 mg/dL or random plasma glucose ≥ 200 mg/dL ; (2) at least 2 instances of elevated plasma glucose concentration (fasting glucose ≥ 140 mg/dL, random plasma glucose ≥ 200 mg/dL, or oral glucose tolerance test ≥ 200 mg/dL after 2 hours) on different occasions in the absence of symptoms; or (3) treatment with insulin or an oral hypoglycemic agent. The criteria for diagnosis of diabetes changed in 1997, when a fasting glucose level of ≥ 126 mg/dL was made the diagnostic threshold. Ascertainment of Gestational Diabetes Women in the NHS II were asked to report the diagnosis of gestational diabetes on each biennial questionnaire. To validate the study's diagnostic criteria, 20,422 participants completed a detailed questionnaire, and 92% corroborated their diagnosis. A review of medical records for 120 of these women	Lifetime lactation history among parous women was stratified into 6 groups: None (referent), > 0 to 3 mo, > 3 to 6 mo, > 6 to 11 mo, > 11 to 23 mo, and > 23 mo. In the NHS, lactation information was used from the 1986 questionnaire. In the NHS II analysis, lactation history was derived from self-reported pregnancies assessed every 2 years and lactation reports in 1997 and 2003. Lifetime duration was updated every 2 years. Linear trend was assessed using midpoints of lactation categories. All models were age-adjusted. Potential confounders, including parity, body mass index at age 18 years, diet, physical activity, family history of diabetes, and smoking status, were included in the multivariate model.	In the NHS cohort, 83,585 parous women reported lifetime duration of lactation; of these, 64% had ever breastfed. In the NHS II cohort, 73,418 parous women reported duration of lactation, and 85% had ever breastfed. In the NHS, women who had ever breastfed had a covariate-adjusted HR for type 2 diabetes of 0.97 (95%CI 0.91 - 1.02) compared with women who never breastfed. There was a modest but statistically significant inverse association between duration of lactation and the risk of type 2 diabetes. In the multivariate-adjusted model including current BMI, each additional year of lactation was associated with an HR of 0.96 (95%CI 0.92 - 0.99) for type 2 diabetes. Among women who had ever breastfed in the NHS II, the covariate-adjusted HR for type 2 diabetes was 0.90 (95%CI 0.77 - 1.04). Each year of lactation was associated with a covariate-adjusted HR of 0.84 (95%CI 0.78 - 0.89). When BMI was added to this model, the HR was 0.88 (95%CI 0.82 - 0.94)	A: strong, B: moderate, C: weak	A	B	C
			Selection	x		
			Study design	x		
			Confounder	x		
			Blinding		x	
			Data collection	x		
			Withdraw and dropout	x		
			Analyses	x		
			Intervention integrity			
Overall: A						

Outcome Definition	Statistical analyses and confounders adjusted	Results	Bias/limitations Comments
confirmed definite or probable gestational diabetes in 94%. To assess screening for gestational diabetes, a random sample of 100 study participants was surveyed, of whom 83% reported undergoing a 50-g, 1-hour glucose challenge test. Gestational diabetes history was not assessed in the NHS cohort.		for each additional year of lactation. Women with a history of GDM had an increased risk of type 2 diabetes in the NHS II cohort, with 624 cases per 100,000 person years compared with 118 cases per 100,000 person-years among those without such a history. Lactation had no effect on diabetes risk in the GDM group, with a covariate-adjusted HR of 0.96 (95% CI 0.84 - 1.09) per additional year of lactation. The effects of exclusive versus total breastfeeding could be compared in the NHS II cohort data. In models controlling for age and parity, each year of lifetime exclusive breastfeeding was associated with an HR for type 2 diabetes of 0.63 (95%CI 0.54 - 0.73), while each year of total breastfeeding was associated with an HR of 0.76 (95%CI 0.71 - 0.81).	

Appendix D. List of Excluded Studies

AbouSaleh MT, Ghubash R, Karim L, et al. Hormonal aspects of postpartum depression. *Psychoneuroendocrinology* 1998 Jul;23(5):465-75.

Developing country

AbuSabha R, Greene G. Body weight, body composition, and energy intake changes in breastfeeding mothers. *J Hum Lact* 1998 Jun;14(2):119-24. **Non-comparative study**

Adair LS, Pollitt E, Mueller WH. Maternal anthropometric changes during pregnancy and lactation in a rural Taiwanese population. *Hum Biol* 1983 Dec;55(4):771-87. **Non-comparative study**

Affinito P, Tommaselli GA, di Carlo C, et al. Changes in bone mineral density and calcium metabolism in breastfeeding women: a one year follow-up study. *J Clin Endocrinol Metab* 1996 Jun;81(6):2314-8. **Less than 1 year follow-up**

Alder E, Bancroft J. The relationship between breast feeding persistence, sexuality and mood in postpartum women. *Psychol Med* 1988 May;18(2):389-96. **N<100**

Alder EM, Cox JL. Breast feeding and post-natal depression. *J Psychosom Res* 1983;27(2):139-44. **N<100**

Alm B, Wennergren G, Norvenius SG, et al. Breast feeding and the sudden infant death syndrome in Scandinavia. *Arch Dis Child* 2002 June; 86(6): 400-2. **Included in MA**

Andreev A, Arjas E. Acute middle ear infection in small children: a Bayesian analysis using multiple time scales. *Lifetime Data Anal* 1998;4(2):121-37. **Not relevant**

Aniansson G, Svensson H, Becker M, et al. Otitis media and feeding with breast milk of children with cleft palate. *Scand J Plast Reconstr Surg Hand Surg* 2002;36(1):9-15. **Otitis media in children with cleft palate**

Atalah E, Lagos I, Grez M, et al. [Effect of lactation on the weight and body composition of wet nurses]. [Spanish]. *Archivos Latinoamericanos de Nutricion* 1983 Sep;33(3):649-63. **Non-English language publication**

Auestad N, Scott DT, Janowsky JS, et al. Visual, cognitive, and language assessments at 39 months: a follow-up study of children fed formulas containing

long-chain polyunsaturated fatty acids to 1 year of age. *Pediatrics* 2003 Sep;112(3 Pt 1):e177-e183. **Not relevant**

Barbosa L, Butte NF, Villalpando S, et al. Maternal energy balance and lactation performance of Mesoamerindians as a function of body mass index. *Am J Clin Nutr* 1997 Sep;66(3):575-83. **Not relevant**

Bastian LA, West NA, Corcoran C, et al. Number of children and the risk of obesity in older women. *Prev Med* 2005 Jan;40(1):99-104. **Case control study**

Batstra L, Neeleman J, Hadders-Algra M. Can breast feeding modify the adverse effects of smoking during pregnancy on the child's cognitive development? *J Epidemiol Community Health* 2003 Jun;57(6):403-4. **Not relevant**

Bauer DC, Browner WS, Cauley JA, et al. Factors associated with appendicular bone mass in older women. The Study of Osteoporotic Fractures Research Group.[see comment]. *Ann Intern Med* 1993 May 1;118(9):657-65. **Cross-sectional study**

Becher H, Schmidt S, ChangClaude J. Reproductive factors and familial predisposition for breast cancer by age 50 years. A case-control-family study for assessing main effects and possible gene-environment interaction.[see comment]. *Int J Epidemiol* 2003 Feb;32(1):38-48. **Included in MA**

Bjorkelund C, Lissner L, Andersson S, et al. Reproductive history in relation to relative weight and fat distribution. *Int J Obes Relat Metab Disord* 1996 Mar;20(3):213-9. **Cross-sectional study**

Boonyaratavej N, Suriyawongpaisal P, Takkinsatien A, et al. Physical activity and risk factors for hip fractures in Thai women. *Osteoporos Int* 2001;12(3):244-8. **Developing country**

Bouwstra H, Dijck-Brouwer DA, Boehm G, et al. Long-chain polyunsaturated fatty acids and neurological developmental outcome at 18 months in healthy term infants. *Acta Paediatr* 2005 Jan;94(1):26-32. **Not relevant**

Bouwstra H, DijckBrouwer DA, Wildeman JA, et al. Long-chain polyunsaturated fatty acids have a positive effect on the quality of general movements of healthy term infants. *Am J Clin Nutr* 2003 Aug;78(2):313-8. **Not relevant**

Bradshaw MK, Pfeiffer S. Feeding mode and anthropometric changes in primiparas. *Hum Biol* 1988 Apr;60(2):251-61. **Insufficient sample size**

Butte NF, Garza C, Stuff JE, et al. Effect of maternal diet and body composition on lactational performance. *Am J Clin Nutr* 1984 Feb;39(2):296-306. **Non-comparative study**

Butte NF, Hopkinson JM. Body composition changes during lactation are highly variable among women. [Review] [45 refs]. *J Nutr* 1998 Feb;128(2:Suppl):Suppl-385S. **Review**

Carranza-Lira S, Mera JP. Influence of number of pregnancies and total breast-feeding time on bone mineral density. *Int J Fertil Womens Med* 2002 Jul;47(4):169-71. **Developing country**

Chan GM, Ronald N, Slater P, et al. Decreased bone mineral status in lactating adolescent mothers. *J Pediatr* 1982 Nov;101(5):767-70. **Less than 1 year follow-up**

Chan SM, Nelson EA, Leung SS, et al. Bone mineral density and calcium metabolism of Hong Kong Chinese postpartum women--a 1-y longitudinal study. *Eur J Clin Nutr* 2005 Jul;59(7):868-76. **Less than 1 year follow-up**

Chang-Claude J, Eby N, Kiechle M, et al. Breastfeeding and breast cancer risk by age 50 among women in Germany. *Cancer Causes Control* 2000 Sep;11(8):687-95. **Pre 2001**

Coitinho DC, Sichieri R, quino Benicio MH. Obesity and weight change related to parity and breast-feeding among parous women in Brazil. *Public Health Nutr* 2001 Aug;4(4):865-70. **Developing country**

Cox JL, Chapman G, Murray D, et al. Validation of the Edinburgh Postnatal Depression Scale (EPDS) in non-postnatal women. *J Affect Disord* 1996 Jul 29;39(3):185-9. **Reference, not relevant**

Cox JL, Connor Y, Kendell RE. Prospective study of the psychiatric disorders of childbirth. *Br J Psychiatry* 1982 Feb;140:111-7. **N<100**

Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale.[see comment]. *Br J Psychiatry* 1987 Jun;150:782-6. **N<100**

Cox ML, Khan SA, Gau DW, et al. Determinants of forearm bone density in premenopausal women: a study in one general practice. *Br J Gen Pract* 1991 May;41(346):194-6. **No BF exposure**

Cramer DW, Welch WR, Scully RE, et al. Ovarian cancer and talc: a case-control study. *Cancer* 1982 Jul 15;50(2):372-6. **Not relevant**

Dalton K. Prospective study into puerperal depression. *Br J Psychiatry* 1971 Jun;118(547):689-92. **N<100**

Dennis KJ, Bytheway WR. Changes in body weight after delivery. *J Obstet Gynaecol Br Commonw* 1965;(72):94-102. **Breastfeeding exclusivity unclear**

Dequeker J, Tobing L, Rutten V, et al. Relative risk factors for osteoporotic fracture: a pilot study of the MEDOS questionnaire. *Clin Rheumatol* 1991 Mar;10(1):49-53. **No exposure; study 2: cross-sectional study**

Dewey KG, Cohen RJ, Brown KH, et al. Effects of exclusive breastfeeding for four versus six months on maternal nutritional status and infant motor development: results of two randomized trials in Honduras. *J Nutr* 2001 Feb;131(2):262-7. **Not relevant**

Duffy SW, Jakes RW, Ng FC, et al. Interaction of dense breast patterns with other breast cancer risk factors in a case-control study. *Br J Cancer* 2004 Jul 19;91(2):233-6. **Population not well-defined**

Dufour DL, Reina JC, Spurr GB. Energy intake and expenditure of free-living, lactating Colombian women in an urban setting. *Eur J Clin Nutr* 2002 Mar;56(3):205-13. **Developing country**

Dugdale AE, Eaton-Evans J. The effect of lactation and other factors on post-partum changes in body-weight and triceps skinfold thickness. *Br J Nutr* 1989 Mar;61(2):149-53. **Unclear exclusive breastfeeding**

Egan KM, Stampfer MJ, Rosner BA, et al. Risk factors for breast cancer in women with a breast cancer family history. *Cancer Epidemiol Biomarkers Prev* 1998 May;7(5):359-64. **Pre 2001**

Engel J, Anteunis L, Volovics A, et al. Risk factors of otitis media with effusion during infancy. *Int J Pediatr Otorhinolaryngol* 1999 May 25;48(3):239-49. **High risk patients included, with 9 ventilated patients**

Enger SM, Ross RK, Paganini-Hill A, et al. Breastfeeding experience and breast cancer risk among postmenopausal women. *Cancer Epidemiol Biomarkers Prev* 1998 May;7(5):365-9. **Pre 2001**

Erkkola M, Pigg HM, Virta-Autio P, et al. Infant feeding patterns in the Finnish type I diabetes prediction and prevention nutrition study cohort. *Eur J Clin Nutr* 2005 Jan;59(1):107-13. **No outcomes of interest**

Findeisen M, Vennemann M, Brinkmann B, et al. German study on sudden infant death (GeSID): design, epidemiological and pathological profile. *Int J of Legal Med* 2004 Jun;118(3):163-9. **Duplicate data of Venneman 2005**

Fituch CC, Palkowetz KH, Goldman AS, et al. Concentrations of IL-10 in preterm human milk and in milk from mothers of infants with necrotizing enterocolitis. *Acta Paediatr* 2004 Nov;93(11):496-500. **Not relevant**

Florey CD, Leech AM, Blackhall A. Infant feeding and mental and motor development at 18 months of age in first born singletons.[see comment]. *Int J Epidemiol* 1995;24 Suppl 1:S21-S26. **Included in systematic review**

Fox KM, Magaziner J, Sherwin R, et al. Reproductive correlates of bone mass in elderly women. Study of Osteoporotic Fractures Research Group. *J Bone Miner Res* 1993 Aug;8(8):901-8. **Cross-sectional study**

Fregene A, Newman LA. Breast cancer in sub-Saharan Africa: how does it relate to breast cancer in African-American women? *Cancer* 2005 Apr 15;103(8):1540-50. **Developing country**

Furberg H, Newman B, Moorman P, et al. Lactation and breast cancer risk. *Int J Epidemiol* 1999 Jun;28(3):396-402. **Pre 2001**

Gale CR, Martyn CN. Breastfeeding, dummy use, and adult intelligence.[see comment]. *Lancet* 1996 Apr 20;347(9008):1072-5. **Included in systematic review**

Galler JR, Ramsey FC, Harrison RH, et al. Postpartum maternal moods and infant size predict performance on a national high school entrance examination. *J Child Psychol Psychiatry* 2004 Sep;45(6):1064-75. **No outcomes of interest**

Gao YT, Shu XO, Dai Q, et al. Association of menstrual and reproductive factors with breast cancer risk: results from the Shanghai Breast Cancer Study. *Int J Cancer* 2000 Jul 15;87(2):295-300. **Developing country**

Geervani P, Jayashree G. A study on nutritional status of adolescent and adult pregnant and lactating women and growth of their infants. *J Trop Pediatr* 1988 Oct;34(5):234-7. **Developing country**

Ghannam NN, Hammami MM, Bakheet SM, et al. Bone mineral density of the spine and femur in healthy Saudi females: relation to vitamin D status, pregnancy, and lactation. *Calcif Tissue Int* 1999 Jul;65(1):23-8. **Cross-sectional study, developing country**

Gilliland FD, Hunt WC, Baumgartner KB, et al. Reproductive risk factors for breast cancer in Hispanic and non-Hispanic white women: the New Mexico Women's Health Study. *Am J Epidemiol* 1998 Oct 1;148(7):683-92. **Pre 2001**

Greene GW, Smiciklas-Wright H, Scholl TO, et al. Postpartum weight change: how much of the weight gained in pregnancy will be lost after delivery? *Obstet Gynecol* 1988 May;71(5):701-7. **Not relevant**

Hadji P, Ziller V, Kalder M, et al. Influence of pregnancy and breast-feeding on quantitative ultrasonometry of bone in postmenopausal women. *Climacteric* 2002 Sep;5(3):277-85. **Cross-sectional study**

Harlow BL, Weiss NS, Roth GJ, et al. Case-control study of borderline ovarian tumors: reproductive history and exposure to exogenous female hormones. *Cancer Res* 1988 Oct 15;48(20):5849-52. **Borderline tumor**

Harrison GA, Boyce AJ, Platt CM, et al. Body composition changes during lactation in a New Guinea population. *Ann Hum Biology* 1975 Oct;2(4):395-8. **Cross-sectional study**

Hart S, Boylan LM, Carroll S, et al. Brief report: breast-fed one-week-olds demonstrate superior neurobehavioral organization. *J Pediatr Psychology* 2003 Dec;28(8):529-34. **Not relevant**

Hartge P, Hoover R, McGowan L, et al. Menopause and ovarian cancer. *Am J Epidemiol* 1988 May;127(5):990-8. **Duplicate study to Harge '89; no additional data**

Hawker GA, Forsmo S, Cadarette SM, et al. Correlates of forearm bone mineral density in young Norwegian women: the Nord-Trondelag Health Study. *Am J Epidemiol* 2002 Sep 1;156(5):418-27. **Cross-sectional study**

Hejar AR, Chong FB, Rosnan H, et al. Breast cancer and lifestyle risks among Chinese women in the Klang Valley in 2001. *Med J Malaysia* 2004 Jun;59(2):226-32. **Developing country**

Hildreth NG, Kelsey JL, Li Volsi VA, et al. An epidemiologic study of epithelial carcinoma of the ovary. *Am J Epidemiol* 1981 Sep;114(3):398-405. **No quantitative data**

Hilson JA, Rasmussen KM, Kjolhede CL. High prepregnant body mass index is associated with poor lactation outcomes among white, rural women independent of psychosocial and demographic correlates. *J Hum Lact* 2004 Feb;20(1):18-29. **No outcomes of interest**

Holt GM, Wolkind SN. Early abandonment of breast feeding: causes and effects. *Child Care Health Dev* 1983 Nov;9(6):349-55. **N<100**

Honda A, Kurabayashi T, Yahata T, et al. Lumbar bone mineral density changes during pregnancy and lactation. *Int J Gynaecol Obstet* 1998 Dec;63(3):253-8. **Cross-sectional study**

Horwood LJ, Fergusson DM. Breastfeeding and later cognitive and academic outcomes. *Pediatrics* 1998 Jan;101(1):E9. **Included in systematic review**

Hu JF, Zhao XH, Chen JS, et al. Bone density and lifestyle characteristics in premenopausal and postmenopausal Chinese women. *Osteoporos Int* 1994 Nov;4(6):288-97. **Cross-sectional study**

Hummel M, Fuchtenbusch M, Schenker M, et al. No major association of breast-feeding, vaccinations, and childhood viral diseases with early islet autoimmunity in the German BABYDIAB Study.[see comment]. *Diabetes Care* 2000 Jul;23(7):969-74. **No outcomes of interest**

Huo D, Lauderdale DS, Li L. Influence of reproductive factors on hip fracture risk in Chinese women. *Osteoporos Int* 2003 Aug;14(8):694-700. **Developing country**

Jacobson SW, Chiodo LM, Jacobson JL. Breastfeeding effects on intelligence quotient in 4- and 11-year-old children. *Pediatrics* 1999 May;103(5):e71. **Pre 2001**

Jacobson SW, Jacobson JL, Frye KF. Incidence and correlates of breast-feeding in socioeconomically disadvantaged women. *Pediatrics* 1991 Oct;88(4):728-36. **N<100**

Johnell O, Nilsson BE. Life-style and bone mineral mass in perimenopausal women. *Calcif Tissue Int* 1984 Jul;36(4):354-6. **Cross-sectional study**

Jones G, Scott FS. A cross-sectional study of smoking and bone mineral density in premenopausal parous women: effect of body mass index, breastfeeding, and sports participation. *J Bone Miner Res* 1999 Sep;14(9):1628-33. **No BF exposure, cross-sectional study**

Kac G, Benicio MH, Velasquez-Melendez G, et al. Breastfeeding and postpartum weight retention in a cohort of Brazilian women. *Am J Clin Nutr* 2004 Mar;79(3):487-93. **Developing country**

Kalkwarf HJ, Specker BL. Bone mineral loss during lactation and recovery after weaning. *Obstet Gynecol* 1995 Jul;86(1):26-32. **Less than 1 year follow-up**

Kamitsuka MD, Horton MK, Williams MA. The incidence of necrotizing enterocolitis after introducing standardized feeding schedules for infants between 1250 and 2500 grams and less than 35 weeks of gestation. *Pediatrics* 2000 Feb;105(2):379-84. **Not relevant**

Keeping JD, Najman JM, Morrison J, et al. A prospective longitudinal study of social, psychological and obstetric factors in pregnancy: response rates and demographic characteristics of the 8556 respondents. *Br J Obstet Gynaecol* 1989 Mar;96(3):289-97. **Not relevant**

Kendell RE, McGuire RJ, Connor Y, et al. Mood changes in the first three weeks after childbirth. *J Affect Disord* 1981 Dec;3(4):317-26. **N<100**

Khedr EM, Farghaly WM, Amry S, et al. Neural maturation of breastfed and formula-fed infants. *Acta Paediatr* 2004 Jun;93(6):734-8. **Not relevant**

Koetting CA, Wardlaw GM. Wrist, spine, and hip bone density in women with variable histories of lactation. *Am J Clin Nutr* 1988 Dec;48(6):1479-81. **Cross-sectional study**

Kramer FM, Stunkard AJ, Marshall KA, et al. Breast-feeding reduces maternal lower-body fat. *J Am Diet Assoc* 1993 Apr;93(4):429-33. **Insufficient sample size**

Laizner AM, Jeans ME. Identification of predictor variables of a postpartum emotional reaction. *Health Care Women Int* 1990;11(2):191-207. **Not relevant**

Lederman SA, Alfasi G, Deckelbaum RJ. Pregnancy-associated obesity in black women in New York City. *Matern Child Health J* 2002 Mar;6(1):37-42. **Not relevant**

Lee CY, Ko IS, Kim HS, et al. Development and validation study of the breast cancer risk appraisal for Korean women. *Nurs Health Sci* 2004 Sep;6(3):201-7. **No hormonal data given**

Lee EO, Ahn SH, You C, et al. Determining the main risk factors and high-risk groups of breast cancer using a predictive model for breast cancer risk assessment in South Korea. *Cancer Nurs* 2004 Sep;27(5):400-6. **Not relevant**

Li R, Jewell S, Grummer-Strawn L. Maternal obesity and breast-feeding practices. *Am J Clin Nutr* 2003 Apr;77(4):931-6. **Not relevant**

Lim S, Joung H, Shin CS, et al. Body composition changes with age have gender-specific impacts on bone mineral density. *Bone* 2004 Sep;35(3):792-8. **No BF exposure, cross-sectional study**

Lovelady CA, Garner KE, Moreno KL, et al. The effect of weight loss in overweight, lactating women on the growth of their infants.[see comment]. *N Engl J Med* 2000 Feb 17;342(7):449-53. **Not relevant**

Lumachi F, Ermani M, Brandes AA, et al. Breast cancer risk in healthy and symptomatic women: results of a multivariate analysis. A case-control study. *Biomed Pharmacother* 2002 Oct;56(8):416-20. **Population not well-defined**

Malloy MH, Berendes H. Does breast-feeding influence intelligence quotients at 9 and 10 years of age? *Early Hum Dev* 1998 Jan 9;50(2):209-17. **Included in systematic review**

Marcus PM, Baird DD, Millikan RC, et al. Adolescent reproductive events and subsequent breast cancer risk. *Am J Public Health* 1999 Aug;89(8):1244-7. **Pre 2001**

Marini A, Vegni C, Gangi S, et al. Influence of different types of post-discharge feeding on somatic growth, cognitive development and their correlation in very low birthweight preterm infants. *Acta Paediatr Suppl* 2003 Sep;91(441):18-33. **Not relevant**

Matsumoto I, Kosha S, Noguchi S, et al. Changes of bone mineral density in pregnant and postpartum women. *J Obstet Gynaecol* 1995 Oct;21(5):419-25. **Less than 1 year follow-up**

McCredie M, Paul C, Skegg DC, et al. Reproductive factors and breast cancer in New Zealand. *Int J Cancer* 1998 Apr 13;76(2):182-8. **Pre 2001**

McCredie M, Paul C, Skegg DC, et al. Breast cancer in Maori and non-Maori women. *Int J Epidemiol* 1999 Apr;28(2):189-95. **Pre 2001**

McKeown T, Aaby P. The influence of reproduction on body weight in women. *J Endocrinol* 1957;15:393-409. **Cross-sectional study**

McLennan JD, Kotelchuck M, Cho H. Prevalence, persistence, and correlates of depressive symptoms in a national sample of mothers of toddlers. *J Am Acad Child Adolesc Psychiatry* 2001 Nov;40(11):1316-23. **Not relevant**

Meeske K, Press M, Patel A, et al. Impact of reproductive factors and lactation on breast carcinoma in situ risk. *Int J Cancer* 2004 May 20;110(1):102-9. **No hormonal data given**

Melton LJ, III, Bryant SC, Wahner HW, et al. Influence of breastfeeding and other reproductive factors on bone mass later in life. *Osteoporos Int* 1993 Mar;3(2):76-83. **Cross-sectional study**

Mezzacappa ES, Katlin ES. Breast-feeding is associated with reduced perceived stress and negative mood in mothers. *Health Psychol* 2002 Mar;21(2):187-93. **N<100**

Mihatsch WA, von Schoenaich P, Fahrenstich H, et al. Randomized, multicenter trial of two different formulas for very early enteral feeding advancement in extremely-low-birth-weight infants.[erratum appears in *J Pediatr Gastroenterol Nutr* 2002 Nov;35(5):739]. *J Pediatr Gastroenterol Nutr* 2001 Aug;33(2):155-9. **Not relevant**

Milner JD, Stein DM, McCarter R, et al. Early infant multivitamin supplementation is associated with increased risk for food allergy and asthma. *Pediatrics* 2004 Jul;114(1):27-32. **Not relevant, 25% premature infants**

Minami Y, Ohuchi N, Taeda Y, et al. Risk factors for benign breast disease according to histopathological type: comparisons with risk factors for breast cancer. *Jpn J Cancer Res* 1998 Feb;89(2):116-23. **Pre 2001**

Minami Y, Tsubono Y, Nishino Y, et al. The increase of female breast cancer incidence in Japan: emergence of birth cohort effect. *Int J Cancer* 2004 Mar 1;108(6):901-6. **Review**

Misri S, Sinclair DA, Kuan AJ. Breast-feeding and postpartum depression: is there a relationship? *Can J Psychiatry* 1997 Dec;42(10):1061-5. **N<100**

Modugno F, Ness RB, Allen GO, et al. Oral contraceptive use, reproductive history, and risk of epithelial ovarian cancer in women with and without endometriosis. *Am J Obstet Gynecol* 2004 Sep;191(3):733-40. **No breastfeeding data**

Monobe H, Ishibashi T, Fujishiro Y, et al. Factors associated with poor outcome in children with acute otitis media. *Acta Otolaryngol (Stockh)* 2003 Jun;123(5):564-8. **Not relevant**

Montella M, De Marco MR, Romeo F, et al. Survival in ovarian cancers treated at National Cancer Institute in Naples Italy 1985-1990. *Eur J Gynaecol Oncol* 1993;14(3):249-54. **No BF exposure**

More C, Bettembuk P, Bhattoa HP, et al. The effects of pregnancy and lactation on bone mineral density.[see comment]. *Osteoporos Int* 2001;12(9):732-7. **Less than 1 year follow-up**

Morley R. Breast feeding and cognitive outcome in children born prematurely. *Adv Exp Med Biol* 2002;503:77-82. **Review**

Morley R, Fewtrell MS, Abbott RA, et al. Neurodevelopment in children born small for

gestational age: a randomized trial of nutrient-enriched versus standard formula and comparison with a reference breastfed group. *Pediatrics* 2004 Mar;113(3 Pt 1):515-21. **Not relevant**

Morrison J, Najman JM, Williams GM, et al. Socio-economic status and pregnancy outcome. An Australian study. *Br J Obstet Gynaecol* 1989 Mar;96(3):298-307. **Not relevant**

Motil KJ, Sheng H, Kertz BL, et al. Lean body mass of well-nourished women is preserved during lactation. *Am J Clin Nutr* 1998 Feb;67(2):292-300. **Exclusive and mixed feeding data combined**

Murphy S, Khaw KT, May H, et al. Parity and bone mineral density in middle-aged women. *Osteoporos Int* 1994 May;4(3):162-16. **Cross-sectional study**

Murray L, Woolgar M, Murray J, et al. Self-exclusion from health care in women at high risk for postpartum depression. *J Public Health Med* 2003 Jun;25(2):131-7. **Not relevant**

Naismith DJ, Ritchie CD. The effect of breast-feeding and artificial feeding on body-weights, skinfold measurements and food intakes of forty-two primiparous women. *Proc Nutr Soc* 1975 Dec;34(3):116A-117A, 1975 Dec. **Abstract**

Najman JM, Morrison J, Williams GM, et al. The employment of mothers and the outcomes of their pregnancies: an Australian study. *Public Health* 1989 May;103(3):189-98. **Not relevant to any outcomes**

Najman JM, Morrison J, Williams GM, et al. Unemployment and reproductive outcome. An Australian study. *Br J Obstet Gynaecol* 1989 Mar;96(3):308-13. **Not relevant to any outcomes**

Nasca PC, Greenwald P, Chorost S, et al. An epidemiologic case-control study of ovarian cancer and reproductive factors. *Am J Epidemiol* 1984 May;119(5):705-13. **No BF exposure**

Newcomb PA, Egan KM, TitusErnstoff L, et al. Lactation in relation to postmenopausal breast cancer. *Am J Epidemiol* 1999 Jul 15;150(2):174-82. **Pre 2001**

Niemela M, Uhari M, Mottonen M. A pacifier increases the risk of recurrent acute otitis media in children in day care centers. *Pediatrics* 1995 Nov;95(5 Pt 1):884-8. **Not relevant**

O'Neill T, Murphy P, Greene VT. Postnatal depression--aetiological factors. *Irish Med J* 1990 Mar;83(1):17-8. **N<100**

Oddy WH. Breastfeeding and asthma in children. A prospective cohort study. *Adv Exp Med Bio*

2000;478:393-4. **Duplicate with Oddy 1999, included in meta-analysis**

Oddy WH, de Klerk NH, Sly PD, et al. The effects of respiratory infections, atopy, and breastfeeding on childhood asthma. *Eur Respir J* 2002 May;19(5):899-905. **Duplicate with Oddy 1999, included in meta-analysis**

Oddy WH, Peat JK, de Klerk NH. Maternal asthma, infant feeding, and the risk of asthma in childhood. *J Allergy Clin Immunol* 2002 Jul;110(1):65-7. **Duplicate with Oddy 1999, included in meta-analysis**

Oddy WH, Sherriff JL, de Klerk NH, et al. The relation of breastfeeding and body mass index to asthma and atopy in children: a prospective cohort study to age 6 years. *Am J Public Health* 2004 Sep;94(9):1531-7. **Duplicate with Oddy 1999, included in meta-analysis**

Olsen LC, Mundt MH. Postpartum weight loss in a nurse-midwifery practice. *J Nurse Midwifery* 1986 Jul;31(4):177-81. **Breastfeeding exclusivity unclear**

Orwoll ES, Bauer DC, Vogt TM, et al. Axial bone mass in older women. Study of Osteoporotic Fractures Research Group. *Ann Intern Med* 1996 Jan 15;124(2):187-96. **Cross-sectional study**

Otto SJ, de Groot RH, Hornstra G. Increased risk of postpartum depressive symptoms is associated with slower normalization after pregnancy of the functional docosahexaenoic acid status. *Prostaglandins Leukotr Essent Fatty Acids* 2003 Oct;69(4):237-43. **Not relevant**

Paine BJ, Makrides M, Gibson RA. Duration of breast-feeding and Bayley's Mental Developmental Index at 1 year of age. *J Paediatr Child Health* 1999 Feb;35(1):82-5. **Pre 2001**

Parker JD. Postpartum weight change. [Review] [21 refs]. *Clin Obstet Gynecol* 1994 Sep;37(3):528-37. **Review**

Paton LM, Alexander JL, Nowson CA, et al. Pregnancy and lactation have no long-term deleterious effect on measures of bone mineral in healthy women: a twin study. *Am J Clin Nutr* 2003 Mar;77(3):707-14. **Not all had 2 year follow-up**

Pearson D, Kaur M, San P, et al. Recovery of pregnancy mediated bone loss during lactation. *Bone* 2004 Mar;34(3):570-8. **Less than 1 year follow-up**

Peters RK, Kjos SL, Xiang A, et al. Long-term diabetogenic effect of single pregnancy in women with previous gestational diabetes mellitus.[see

comment]. *Lancet* 1996 Jan 27;347(8996):227-30. **No breastfeeding exposure**

Pettitt DJ, Knowler WC. Long-term effects of the intrauterine environment, birth weight, and breast-feeding in Pima Indians. *Diabetes Care* 1998 Aug;21(Suppl 2):B138-B141. **Included in systematic review**

Piers LS, Diggavi SN, Thangam S, et al. Changes in energy expenditure, anthropometry, and energy intake during the course of pregnancy and lactation in well-nourished Indian women. *Am J Clin Nutr* 1995 Mar;61(3):501-13. **Developing country**

Pitt B. "Atypical" depression following childbirth. *Br J Psychiatry* 1968 Nov;114(516):1325-35. **N<100**

Plagemann A, Harder T, Franke K, et al. Long-term impact of neonatal breast-feeding on body weight and glucose tolerance in children of diabetic mothers. *Diabetes Care* 2002 Jan;25(1):16-22. **Included in systematic review**

Plagemann A, Harder T, Kohlhoff R, et al. Impact of early neonatal breast-feeding on psychomotor and neuropsychological development in children of diabetic mothers. *Diabetes Care* 2005 Mar;28(3):573-8. **Not relevant**

Potter S, Hannum S, McFarlin B, et al. Does infant feeding method influence maternal postpartum weight loss? *J Am Diet Assoc* 1991 Apr;91(4):441-6. **Retrospective study**

Prentice AM, Whitehead RG, Roberts SB, et al. Long-term energy balance in child-bearing Gambian women. *Am J Clin Nutr* 1981 Dec;34(12):2790-9. **Developing country**

Purdie D, Green A, Bain C, et al. Reproductive and other factors and risk of epithelial ovarian cancer: an Australian case-control study. *Survey of Women's Health Study Group. Int J Cancer* 1995 Sep 15;62(6):678-84. **Duplicate study to Siskind 1997, no additional data**

Purwanto H, Sadjimin T, Dwiprahasto I. Lactation and the risk of breast cancer. *Gan to Kagaku Ryoho [Jpn J Cancer Chemotherapy]* 2000 May;27(Suppl 2):474-81. **Pre 2001**

Rami B, Schneider U, Imhof A, et al. Risk factors for type I diabetes mellitus in children in Austria. *Eur J Pediatr* 1999 May;158(5):362-6. **Duplicate data – included in EURODIAB**

Richards M, Wadsworth M, Rahimi-Foroushani A, et al. Infant nutrition and cognitive development in the first offspring of a national UK birth cohort. *Dev Med*

Child Neurol 1998 Mar;40(3):163-7. **Included in systematic review**

Rojahn A, Lindgren CG. Enteral feeding in infants <1250 g starting within 24 h post-partum. *Eur J of Pediatr* 2001 Oct;160(10):629-32. **Not relevant**

Rookus MA, Rokebrand P, Burema J, et al. The effect of pregnancy on the body mass index 9 months postpartum in 49 women. *Int J Obes* 1987;11(6):609-18. **Breastfeeding exclusivity unclear**

Rooney BL, Schauburger CW. Excess pregnancy weight gain and long-term obesity: one decade later. *Obstet Gynecol* 2002 Aug;100(2):245-52. **Breastfeeding exclusivity unclear**

Rosenblatt KA, Thomas DB. Lactation and the risk of epithelial ovarian cancer. The WHO Collaborative Study of Neoplasia and Steroid Contraceptives. *Int J Epidemiol* 1993 Apr;22(2):192-7. **Included developing countries**

Rossner S. Weight gain in pregnancy. [Review] [17 refs]. *Human Reproduct* 1997 Oct;12:Suppl-5. **Review, duplicate with Ohlin study**

Rush D, Lumey LH, Ravelli AC, et al. The indirect association of lactation with subsequent perimenopausal body weight. *Eur J of Clin Nutr* 1996 Jan;50(1):12-6. **Cross-sectional study**

Sadurskis A, Kabir N, Wager J, et al. Energy metabolism, body composition, and milk production in healthy Swedish women during lactation. *Am J Clin Nutr* 1988 Jul;48(1):44-9. **Non-comparative study**

Samuelsson U, Ludvigsson J. Seasonal variation of birth month and breastfeeding in children with diabetes mellitus. *J Pediatr Endocrinol* 2001 Jan;14(1):43-6. **Same subjects as Samuelsson, 1993 in the MA but include results from new analyses**

Schauburger CW, Rooney BL, Brimer LM. Factors that influence weight loss in the puerperium. *Obstet Gynecol* 1992 Mar;79(3):424-9. **No data**

Schneider AP. Risk factor for ovarian cancer. *N Engl J Med* 1987 Aug 20;317(8):508-9. **Abstract**

Schwartz SM, Weiss NS, Daling JR, et al. Incidence of histologic types of uterine sarcoma in relation to menstrual and reproductive history. *Int J Cancer* 1991 Sep 30;49(3):362-7. **No outcomes of interest**

Sheikh GN. Observations of maternal weight behavior during the puerperium. *Am J Obstet Gynecol* 1971 Sep 15;111(2):244-50. **Unclear exclusive breastfeeding**

Singh Y, Sayami P, Sayami G, et al. Nepalese breast cancer in relation to reproductive factors: comparison between Nepalese and Japanese cases. *Anticancer Res* 2002 Jan;22(1A):319-23. **Developing country**

Sinigaglia L, Varenna M, Binelli L, et al. Effect of lactation on postmenopausal bone mineral density of the lumbar spine. *J Reprod Med* 1996 Jun;41(6):439-43. **Cross-sectional study**

Soltesz G, Jeges S, Dahlquist G. Non-genetic risk determinants for type 1 (insulin-dependent) diabetes mellitus in childhood. Hungarian Childhood Diabetes Epidemiology Study Group. *Acta Paediatr* 1994 Jul;83(7):730-5. **Developing country**

Sowers M, Corton G, Shapiro B, et al. Changes in bone density with lactation.[see comment]. *JAMA* 1993 Jun 23;269(24):3130-5. **Less than 1 year follow-up**

Sowers M, Wallace RB, Lemke JH. Correlates of forearm bone mass among women during maximal bone mineralization. *Prev Med* 1985 Sep;14(5):585-96. **Cross-sectional study**

Sowers M, Zhang D, Janney CA. Interpregnancy weight retention patterning in women who breastfed. *J Matern Fetal Med* 1998 Mar;7(2):89-94. **No data**

Stene LC, Joner G, Norwegian Childhood Diabetes Study Group. Atopic disorders and risk of childhood-onset type 1 diabetes in individuals. *Clin Exp Allergy* 2004 Feb;34(2):201-6. **No BF exposure**

Steyn NP, Mann J, Bennett PH, et al. Diet, nutrition and the prevention of type 2 diabetes. [Review] [248 refs]. *Public Health Nutr* 2004 Feb;7(1A):147-65. **Review**

Susman VL, Katz JL. Weaning and depression: another postpartum complication. *Am J Psychiatry* 1988 Apr;145(4):498-501. **Case report**

Tamminen T. The impact of mother's depression on her nursing experiences and attitudes during breastfeeding. *Acta Paediatr Scand Suppl No 344* 1988;77:87-94. **N<100**

Tappin D, Ecob R, Brooke H. Bedsharing, roomsharing, and sudden infant death syndrome in Scotland: a case-control study.[see comment]. *J Pediatr* 2005 Jul;147(1):32-7. **No BF data**

Tarkka MT, Paunonen M, Laippala P. Factors related to successful breast feeding by first-time mothers when the child is 3 months old. *J Adv Nurs* 1999 Jan;29(1):113-8. **Not relevant**

Thomson AM, Billewicz WZ, Thompson B, et al. Body weight changes during pregnancy and lactation in rural African (Gambian) women. *J Obstet Gynaecol Br Commonw* 1966 Oct;73(5):724-33. **Developing country**

Titus-Ernstoff L, Egan KM, Newcomb PA, et al. Exposure to breast milk in infancy and adult breast cancer risk. *J Natl Cancer Inst* 1998 Jun 17;90(12):921-4. **Breast fed infant related cancer**

Tryggvadottir L, Olafsdottir EJ, Gudlaugsdottir S, et al. BRCA2 mutation carriers, reproductive factors and breast cancer risk. *Breast Cancer Res* 2003;5(5):R121-R128. **Included in MA**

Tryggvadottir L, Tulinius H, Eyfjord JE, et al. Breast cancer risk factors and age at diagnosis: an Icelandic cohort study. *Int J Cancer* 2002 Apr 1;98(4):604-8. **Pre 2001**

Tuppurainen M, Kroger H, Honkanen R, et al. Risks of perimenopausal fractures--a prospective population-based study. *Acta Obstet Gynecol Scand* 1995 Sep;74(8):624-8. **No BF exposure**

Tuppurainen M, Kroger H, Saarikoski S, et al. The effect of gynecological risk factors on lumbar and femoral bone mineral density in peri- and postmenopausal women. *Maturitas* 1995 Feb;21(2):137-45. **Cross-sectional study**

Unay B, Sarici SU, Ulas UH, et al. Nutritional effects on auditory brainstem maturation in healthy term infants. *Arch Dis in Child Fetal Neonatal Ed* 2004 Mar;Edition 89(2):F177-F179. **Not relevant**

Ursin G, Bernstein L, Lord SJ, et al. Reproductive factors and subtypes of breast cancer defined by hormone receptor and histology. *Br J Cancer* 2005 Aug 8;93(3):364-71. **No hormonal data given**

Ursin G, Bernstein L, Wang Y, et al. Reproductive factors and risk of breast carcinoma in a study of white and African-American women. *Cancer* 2004 Jul 15;101(2):353-62. **No hormonal data given**

van Hemert AM, Vandenbroucke JP, Birkenhager JC, et al. Prediction of osteoporotic fractures in the general population by a fracture risk score. A 9-year follow-up among middle-aged women. *Am J Epidemiol* 1990 Jul;132(1):123-35. **No outcomes of interest**

Van Raaij JM, Schonk CM, Vermaat-Miedema SH, et al. Energy cost of lactation, and energy balances of well-nourished Dutch lactating women: reappraisal of the extra energy requirements of lactation. *Am J Clin Nutr* 1991 Mar;53(3):612-9. **Non-comparative study**

Vennemann MM, Bajanowski T, Butterfass-Bahloul T, et al. Do risk factors differ between explained sudden unexpected death in infancy (SUDI) and SIDS? Arch Dis Child Epub ahead of print doi:10.1136/adc.2006.101337.2006 Aug 25. **Duplicate of Vennemann 2005**

Verge CF, Howard NJ, Irwig L, et al. Environmental factors in childhood IDDM. A population-based, case-control study. Diabetes Care 1994 Dec;17(12):1381-9. **Included in MA**

Virtanen SM, Rasanen L, Ylonen K, et al. Early introduction of dairy products associated with increased risk of IDDM in Finnish children. The Childhood in Diabetes in Finland Study Group. Diabetes 1993 Dec;42(12):1786-90. **Studies in the MA**

Walker LO, Freeland-Graves J. Lifestyle factors related to postpartum weight gain and body image in bottle- and breastfeeding women. J Obstet Gynecol Neonatal Nurs 1998 Mar;27(2):151-60. **Cross-sectional study**

Webster J, Moore K, McMullan A. Breastfeeding outcomes for women with insulin dependent diabetes. J Hum Lact 1995 Sep;11(3):195-200. **Not relevant**

Weiss NS, Lyon JL, Liff JM, et al. Incidence of ovarian cancer in relation to the use of oral contraceptives. Int J Cancer 1981 Dec;28(6):669-71. **No BF exposure**

Whiteman MK, Hillis SD, Curtis KM, et al. Reproductive history and mortality after breast cancer diagnosis. Obstet Gynecol 2004 Jul;104(1):146-54. **No outcomes of interest**

Winkvist A, Rasmussen KM. Impact of lactation on maternal body weight and body composition. [Review] [71 refs]. J Mammary Gland Biol Neoplasms 1999 Jul;4(3):309-18. **Review**

Wosje KS, Kalkwarf HJ. Lactation, weaning, and calcium supplementation: effects on body composition in postpartum women. Am J Clin Nutr 2004 Aug;80(2):423-9. **Breastfeeding exclusivity unclear**

Yasumizu T, Nakamura Y, Hoshi K, et al. Bone metabolism after human parturition and the effect of lactation: longitudinal analysis of serum bone-related proteins and bone mineral content of the lumbar spine. Endocr J 1998 Oct;45(5):679-86. **Less than 1 year follow-up**

Yen ML, Yen BL, Bai CH, et al. Risk factors for ovarian cancer in Taiwan: a case-control study in a low-incidence population. Gynecol Oncol 2003 May;89(2):318-24. **Developing country**

Young TK, Martens PJ, Taback SP, et al. Type 2 diabetes mellitus in children: prenatal and early infancy risk factors among native Canadians. Arch Pediatr Adolesc Med 2002 Jul;156(7):651-5. **Included in systematic review**

Zhang J, Liu G, Wu J, et al. Pregnancy-associated breast cancer: a case control and long-term follow-up study in China. J Exp Clin Cancer Res 2003 Mar;22(1):23-7. **Developing country**

Zhang M, Lee AH, Binns CW. Reproductive and dietary risk factors for epithelial ovarian cancer in China. Gynecol Oncol 2004 Jan;92(1):320-6. **Developing country**

Zhang M, Xie X, Lee AH, et al. Prolonged lactation reduces ovarian cancer risk in Chinese women. Eur J Cancer Prev 2004 Dec;13(6):499-502. **Developing country**

Zheng T, Duan L, Liu Y, et al. Lactation reduces breast cancer risk in Shandong Province, China. Am J Epidemiol 2000 Dec 15;152(12):1129-35. **No hormonal data given**

Zheng T, Holford TR, Mayne ST, et al. Lactation and breast cancer risk: a case-control study in Connecticut. Br J Cancer 2001 Jun 1;84(11):1472-6. **Pre 2001**

Ziegler AG, Schmid S, Huber D, et al. Early infant feeding and risk of developing type 1 diabetes-associated autoantibodies.[see comment]. JAMA 2003 Oct 1;290(13):1721-8. **No outcomes of interest**

Appendix E. Peer Reviewers

The peer reviewer comments on a preliminary draft of this report were considered by the EPC in preparation of this final report. Synthesis of the scientific literature presented here does not necessarily represent the views of individual reviewers. The authors gratefully acknowledge the peer reviewers.

David Atkins, MD, MPH
Chief Medical Officer
Center for Outcomes and Evidence
Agency for Healthcare Research and Quality
Rockville, Maryland

Pamela D. Hill, PhD, RN, CBE, FAAN
Professor
College of Nursing
University of Illinois at Chicago
Director, Quad Cities Regional Program
Moline, Illinois

Lawrence M. Gartner, MD
Professor Emeritus
Departments of Pediatrics and
Obstetrics/Gynecology
The University of Chicago
Chicago, Illinois

International Formula Council
Atlanta, Georgia

Peer Reviewers

Russell Merritt, MD, PhD, FAAP
Carol Lynn Berseth, MD, FAAP
Jose Saavedra, MD, FAAP
John Wallingford, PhD

Representative for the
American Academy of Pediatrics

Jeanne-Marie Guise, MD, MPH
Associate Professor
Director of Clinical Research
Division of General Obstetrics & Gynecology
Oregon Health and Science University
Portland, Oregon

Nanette Pazdernik, PhD
Center for Breastfeeding Information
La Leche League International
Schaumburg, Illinois

Jane Heinig, PhD, IBCLC
Academic Administrator
Department of Nutrition
Director, Human Lactation Center
University of California, Davis

Laurie P. Shornick, PhD
Assistant Professor
Department of Biology
Saint Louis University
St. Louis, Missouri

...

Long-term effects of breastfeeding

...

A SYSTEMATIC REVIEW

...

Bernardo L. Horta, MD, PhD
Universidade Federal de Pelotas, Pelotas, Brazil

Cesar G. Victora, MD, PhD
Universidade Federal de Pelotas, Pelotas, Brazil



**World Health
Organization**

WHO Library Cataloguing-in-Publication Data

Long-term effects of breastfeeding: a systematic review.

1.Breast feeding. 2.Infant mortality. 3.Infection – prevention and control. 4.Chronic disease – prevention and control. 5.Milk, Human. 6.Review. I.Horta, Bernardo L. II.Victora, Cesar G. III.World Health Organization.

ISBN 978 92 4 150530 7

(NLM classification: WS 125)

© World Health Organization 2013

All rights reserved. Publications of the World Health Organization are available on the WHO web site (www.who.int) or can be purchased from WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel.: +41 22 791 3264; fax: +41 22 791 4857; e-mail: bookorders@who.int). Requests for permission to reproduce or translate WHO publications – whether for sale or for non-commercial distribution – should be addressed to WHO Press through the WHO web site (www.who.int/about/licensing/copyright_form/en/index.html).

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by the World Health Organization to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall the World Health Organization be liable for damages arising from its use.

The named authors alone are responsible for the views expressed in this publication.

Contents

Chapter 1.	Introduction	1
Chapter 2.	Methodological issues	3
Chapter 3.	Search methods	8
Chapter 4.	Review methods	10
Chapter 5.	Overweight and obesity	13
Chapter 6.	Blood pressure	28
Chapter 7.	Total cholesterol	41
Chapter 8.	Type-2 diabetes	51
Chapter 9.	Performance in intelligence tests	57
Chapter 10.	Conclusions	67

Introduction

Breastfeeding has well-established short-term benefits, particularly the reduction of morbidity and mortality due to infectious diseases in childhood. A pooled analysis of studies carried out in middle/low income countries showed that breastfeeding substantially lowers the risk of death from infectious diseases in the first two years of life (1).

Based on data from the United Kingdom Millennium Cohort, Quigley et al (2) estimated that optimal breastfeeding practices could prevent a substantial proportion of hospital admissions due to diarrhea and lower respiratory tract infection. A systematic review by Kramer et al (3) confirmed that exclusive breastfeeding in the first 6 months decreases morbidity from gastrointestinal and allergic diseases, without any negative effects on growth. Given such evidence, it has been recommended that in the first six months of life, every child should be exclusively breastfed, with partial breastfeeding continued until two years of age (4).

Building upon the strong evidence on the short-term effects of breastfeeding, the present review addresses its long-term consequences. Current evidence, mostly from high income countries, suggests that occurrence of non-communicable diseases may be programmed by exposures occurring during gestation or in the first years of life (5–7). Early diets, including the type of milk received, is one of the key exposures that may influence the development of adult diseases.

In 2007, we carried out a systematic review and meta-analysis on the long-term consequences of breastfeeding. The Department of Maternal, Newborn, Child and Adolescent Health of the World Health Organization has now commissioned an update of this review. The following long-term outcomes were reviewed: blood pressure, type-2 diabetes, serum cholesterol, overweight and obesity, and intellectual performance. These outcomes are of great interest to researchers, as made evident by the number of publications identified: 60 new publications were identified since 2006. This report describes the methods, results and conclusions of this updated review.

References

- 1 Effect of breastfeeding on infant and child mortality due to infectious diseases in less developed countries: a pooled analysis. WHO Collaborative Study Team on the Role of Breastfeeding on the Prevention of Infant Mortality. *Lancet* 2000; 355: 451–5.
- 2 Quigley MA, Kelly YJ, Sacker A. Breastfeeding and hospitalization for diarrheal and respiratory infection in the United Kingdom Millennium Cohort Study. *Pediatrics* 2007; 119: e837–42.
- 3 Kramer MS, Kakuma R. The optimal duration of exclusive breastfeeding: a systematic review. *Adv Exp Med Biol* 2004; 554: 63–77.
- 4 World Health Organization, Global strategy for infant and young child feeding. The optimal duration of exclusive breastfeeding, in, Geneva, World Health Organization, 2001.
- 5 Horta BL, Barros FC, Victora CG, et al. Early and late growth and blood pressure in adolescence. *J Epidemiol Community Health* 2003; 57: 226–30.

- 6 Adair L, Dahly D. Developmental determinants of blood pressure in adults. *Annu Rev Nutr* 2005; 25: 407–34.
- 7 Owen CG, Martin RM, Whincup PH, et al. Does breastfeeding influence risk of type 2 diabetes in later life? A quantitative analysis of published evidence. *Am J Clin Nutr* 2006; 84: 1043–54.

Methodological issues

Randomized controlled trials, if properly designed and conducted, provide the best evidence on a causal association between an exposure – such as breastfeeding – and a health or developmental outcome. Randomization increases the likelihood that results will not be affected by confounding (1). Additionally, existing guidelines propose standards for conducting, analyzing and reporting clinical trials, which helps increase the validity of the evidence (2).

On the other hand, the recognition of the short-term benefits of breastfeeding, briefly described in **Chapter 1**, constitutes an ethical challenge to the design of randomized trials aimed at assessing its long-term consequences. Currently, it would be considered unethical to randomly allocate subjects not to receive breastmilk. In contrast, in the early 1980s the evidence on the short-term benefits of breastfeeding was not so clear-cut. At that time, preterm infants admitted to neonatal units could be ethically allocated at random to receive banked breastmilk or formula. Follow-up of these subjects in adolescence has provided experimental evidence on the long-term effects of breastfeeding (3–5).

An alternative to individual randomization to breastfeeding is to allocate groups of mothers to receive – or not to receive – breastfeeding counseling. In Belarus, the Promotion of Breastfeeding Trial (6) randomly assigned maternity hospitals and their affiliated polyclinics to the Baby-Friendly Hospital Initiative. The proportion of infants exclusively breastfed at 3 and 6 months was substantially higher among infants from the intervention group. This trial is ethically sound because mothers were randomly assigned to receive intense breastfeeding promotion, compared to usual care in the hospitals. The follow-up of this study has provided invaluable evidence on the long-term consequence of breastfeeding (7–8). On the other hand, compliance to the intervention was far from universal, only 43.3% of the infants in the intervention group were exclusively breastfed at 3 months compared to 6.4% in the comparison arm, and therefore both groups represent a mixture of breastfeeding practices. As a consequence, the effect of breastfeeding itself on outcomes is underestimated, and statistical power is reduced.

Because of the small number of randomized controlled trials with sufficient follow-up time, most of the evidence on the long-term effects of breastfeeding is derived from observational studies. Prospective birth cohort studies are the next-best design in terms of strength of evidence.

Below, we discuss the strengths and weaknesses of observational studies, as well as approaches that may help overcome their main shortcomings.

Factors affecting internal validity

Losses to follow-up

If losses to follow-up are high, selection bias may be introduced. This may affect both randomized and observational studies. In order to assess the study susceptibility to selection bias, baseline data,

such as breastfeeding duration, should be compared between those subjects who were followed up and those who were not. If attrition rates are not related to breastfeeding duration or other baseline characteristics, selection bias is unlikely (9). However, not every study provides such information.

Misclassification

When assessment of exposure or the outcomes is inaccurate, misclassification may occur. Misclassification may be differential or non-differential.

Retrospective studies are more susceptible to recall bias and direction of bias may change. For example, Huttly et al (10) observed that Brazilian mothers of high socioeconomic status tended to overestimate the breastfeeding duration, whereas among poor mothers this was not the case. This differential recall of breastfeeding duration would tend to overestimate the benefits of breastfeeding because high socioeconomic status is associated with a lower risk of chronic non transmissible diseases.

On the other hand, if the measurement error is not related to exposure or outcome, non-differential misclassification occurs. Such bias underestimates the measure of association, and, therefore, reduces the likelihood of reporting a significant association. Indeed, in a meta-analysis on the relationship between maternal smoking in pregnancy and breastfeeding duration, the odds ratio for weaning at 3 months was inversely related to the length of recall for exposure and outcome (11).

Unfortunately, very few of the studies included in this review address these issues. We attempted to address it by stratifying studies according to the length of recall of breastfeeding information, but admittedly this is only a proxy for misclassification.

Confounding

Confounding is one of the challenges in interpreting the evidence of observational studies. Even large studies that managed to measure the possible confounders may still be affected by residual confounding, if the confounder variables were not properly measured or adjusted for. Some methods have been suggested to improve causal inference. These include comparison of siblings in within-family analyses, which allow controlling for unmeasured maternal and family variables (socioeconomic status, maternal variables) as well as for self-selection bias, because these characteristics are shared among siblings. Usually, sibling studies assess the effect of discordance on breastfeeding duration or complementary feeding on the outcome. Gillman et al (12) used this design to investigate the association between breastfeeding and overweight (see [Chapter 5](#)). A limitation of these studies, is that heterogeneity in breastfeeding duration is smaller among siblings than that observed among unrelated individuals and the sample size for the sibling analysis are smaller, decreasing statistical power.

Another strategy involves the comparison of observational studies with a different confounding structure. In this approach, if an association is causal, the association should be observed in every setting, in spite of differing confounding structures. Brion et al (13) compared the effect of breastfeeding on blood pressure, body mass index and intelligence quotients in two cohorts, one in the United Kingdom (ALSPAC) where breastfeeding duration is positively associated with family income, and another in Brazil (Pelotas) where there is no such association. In ALSPAC, even after controlling for confounding for socioeconomic status, breastfeeding was inversely related to blood pressure and body mass index and positively with performance in intelligence tests. On the other hand, in Pelotas, breastfeeding was only associated with higher performance in intelligence tests. Therefore, the observed effect of breastfeeding on blood pressure and body mass index may be due to residual confounding,

whereas the higher performance in intelligence tests among subjects who were breastfed is likely to be causal. This approach should be further used; replication of an association across different settings with heterogeneous confounding structure would improve causal inference in observational studies.

Self-selection

In high-income settings, breastfeeding mothers are more likely to be health-conscious, and, therefore, also more likely to promote healthy habits to their infants, including prevention of obesity, promotion of physical activity and intellectual stimulation. For example, cognitive stimulation and emotional support is higher among children who are breastfed (14). Because stimulation and emotional support are positively associated with performance in intelligence tests and cognition, studies assessing the long-term consequences of breastfeeding on performance in intelligence tests should adjust their estimates to home stimulation. In this situation, the self-selection bias is measured by proxy, and treated as a confounder.

Even when it is not possible to adjust for proxies for self-selection, this possibility should be considered when interpreting the results.

Adjustment for potential mediating factors

Several studies on the long-term consequences of breastfeeding present estimates of effect that are adjusted for variables that may represent pathways in the causal chain leading from breast feeding to the outcomes being studied. In this review, the most common example was the inclusion of adult weight measures when studying outcomes such as diabetes, hypertension or cholesterol levels, which themselves are influenced by weight.

Adjustment for such mediating factors will tend to underestimate the overall effect of breastfeeding, and the adjusted estimate will only represent the part of the effect that is not mediated by current weight (15).

Main sources of heterogeneity among studies

Heterogeneity among studies is unavoidable, and the question is not whether heterogeneity is present, but if it seriously undermines the conclusions. Well-conducted meta-analyses should incorporate a detailed investigation of potential sources of heterogeneity (16). The following possible sources of heterogeneity were considered for all reviews in the present meta-analyses.

Year of birth

Studies on the long-term consequence of breastfeeding have included subjects born at different times in the past. During this period, the diets of non-breastfed infants in high-income countries have changed markedly. In the first decades of the 20th century, most non-breastfed infants received formulations based on whole cow's milk or top milk (17), with high sodium concentrations and levels of cholesterol and fatty acids that were similar to those in mature breastmilk. By the 1950s, commercially prepared formulas became increasingly popular. At this time, formulas tended to have high sodium concentrations and low levels of iron and essential fatty acids. Starting in the 1980s, sodium content was reduced and nowadays the majority of formulas have levels that are similar to those in breastmilk (18). Therefore, the year of birth of the studied population may affect the long-term effects of breastfeeding, representing a source of heterogeneity among studies. This possibility was investigated in the present review.

Length of recall of breastfeeding

Misclassification of breastfeeding duration has been discussed above. Feeding histories were often assessed retrospectively, and the length of recall has varied widely among studies. As previously mentioned, length of recall is related to misclassification of breastfeeding duration. Some studies suggest that bias tends to increase with the time elapsed since weaning, with mothers who breastfed for a short period being more likely to exaggerate breastfeeding duration, while the opposite is observed for women who breastfed for long periods (10,19). Therefore, length of recall is a potential source of heterogeneity among studies.

Source of information on breastfeeding duration

Studies on the long-term effect of breastfeeding have usually assessed infant feeding by maternal recall, while others relied on information collected by health workers, or on the subjects' own reports. Marmot et al (20) reported that misclassification of the subjects' own infant feeding is higher among subjects who were bottle-fed. The direction of such bias will depend whether or not classification error is associated with factors related to morbidity in adulthood.

Categories of breastfeeding duration

Studies on the long-term effect of breastfeeding have compared different groups according to breastfeeding duration. Some studies compared ever-breastfed subjects to those never breastfed, whereas other studies compared subjects breastfed for less than a given number of months to those breastfed for longer periods. The comparison of ever versus never breastfed makes sense if the first hours of life are considered as a critical window for the programming effect of breastfeeding, for example if an epigenetic mechanism is being postulated (21). On the other hand, if there is no critical-window effect, but rather a cumulative effect of breastfeeding, studies that compared ever vs. never breastfed subjects will tend to underestimate any association. The classification of breastfeeding duration is another factor to be considered in heterogeneity analyses.

Study setting

Most of the studies on the long-term consequences of breastfeeding have been carried out in high-income countries. The findings from these studies may not hold for populations exposed to different environmental and nutritional factors because of differences in the type of milk fed to non-breastfed infants. In high-income countries, the babies usually receive industrialized formula, whereas many non-breastfed infants in low and middle income countries receive whole or diluted animal milk.

Most exposures in epidemiological studies (for example, smoking, alcohol drinking, environmental pollutants or specific dietary items) entail a comparison with a group that is not exposed to these risk or protective factors. In contrast, young infants who are not breastfed subjects – the “unexposed” – must receive some type of feeding, and are thus exposed to a variety of foodstuffs, such as industrialized formula, animal milk or traditional weaning foods. The heterogeneity of the unexposed group in breastfeeding studies must be taken into account. This issue is related to the age of the cohort, discussed above, and to the setting of the study, e.g. high or low-income country.

Whenever possible, in light of the information provided by the studies' authors, we tried to assess the role of these variables in the interpretation of the present review.

References

- 1 Chalmers I. Unbiased, relevant, and reliable assessments in health care: important progress during the past century, but plenty of scope for doing better. *BMJ* 1998; 317: 1167–8.
- 2 Schulz KF, Altman DG, Moher D, et al. CONSORT 2010 changes and testing blindness in RCTs. *Lancet*; 375: 1144–6.
- 3 Lucas A, Morley R. Does early nutrition in infants born before term programme later blood pressure? *BMJ* 1994; 309: 304–8.
- 4 Lucas A, Morley R, Cole TJ, et al. Breast milk and subsequent intelligence quotient in children born preterm. *Lancet* 1992; 339: 261–4.
- 5 Singhal A, Cole TJ, Fewtrell M, et al. Breastmilk feeding and lipoprotein profile in adolescents born preterm: follow-up of a prospective randomised study. *Lancet* 2004; 363: 1571–8.
- 6 Kramer MS, Chalmers B, Hodnett ED, et al. Promotion of Breastfeeding Intervention Trial (PROBIT): a randomized trial in the Republic of Belarus. *JAMA* 2001; 285: 413–20.
- 7 Kramer MS, Aboud F, Mironova E, et al. Breastfeeding and child cognitive development: new evidence from a large randomized trial. *Arch Gen Psychiatry* 2008; 65: 578–84.
- 8 Kramer MS, Matush L, Vanilovich I, et al. Effects of prolonged and exclusive breastfeeding on child height, weight, adiposity, and blood pressure at age 6.5 y: evidence from a large randomized trial. *Am J Clin Nutr* 2007; 86: 1717–21.
- 9 Victora CG, Barros FC, Lima RC, et al. The Pelotas birth cohort study, Rio Grande do Sul, Brazil, 1982–2001. *Cad Saude Publica* 2003; 19: 1241–56.
- 10 Huttly SR, Barros FC, Victora CG, et al. Do mothers overestimate breast feeding duration? An example of recall bias from a study in southern Brazil. *Am J Epidemiol* 1990; 132: 572–5.
- 11 Horta BL, Kramer MS, Platt RW. Maternal smoking and the risk of early weaning: a meta-analysis. *Am J Public Health* 2001; 91: 304–7.
- 12 Gillman MW, Rifas-Shiman SL, Berkey CS, et al. Breast-feeding and overweight in adolescence: within-family analysis [corrected]. *Epidemiology* 2006; 17: 112–4.
- 13 Brion MJ, Lawlor DA, Matijasevich A, et al. What are the causal effects of breastfeeding on IQ, obesity and blood pressure? Evidence from comparing high-income with middle-income cohorts. *Int J Epidemiol*.
- 14 Der G, Batty GD, Deary IJ. Effect of breast feeding on intelligence in children: prospective study, sibling pairs analysis, and meta-analysis. *BMJ* 2006; 333: 945.
- 15 Victora CG, Huttly SR, Fuchs SC, et al. The role of conceptual frameworks in epidemiological analysis: a hierarchical approach. *Int J Epidemiol* 1997; 26: 224–7.
- 16 Thompson SG. Why sources of heterogeneity in meta-analysis should be investigated. *BMJ* 1994; 309: 1351–5.
- 17 Barr SI, McCarron DA, Heaney RP, et al. Effects of increased consumption of fluid milk on energy and nutrient intake, body weight, and cardiovascular risk factors in healthy older adults. *J Am Diet Assoc* 2000; 100: 810–7.
- 18 Fomon S. Infant feeding in the 20th century: formula and beikost. *J Nutr* 2001; 131: 409S–20S.
- 19 Promislow JH, Gladen BC, Sandler DP. Maternal recall of breastfeeding duration by elderly women. *Am J Epidemiol* 2005; 161: 289–96.
- 20 Marmot MG, Page CM, Atkins E, et al. Effect of breast-feeding on plasma cholesterol and weight in young adults. *J Epidemiol Community Health* 1980; 34: 164–7.
- 21 Kuh D, Ben-Shlomo Y, Lynch J, et al. Life course epidemiology. *J Epidemiol Community Health* 2003; 57: 778–83.

Search methods

The search methods were exactly the same as in the previous review we published in 2007. Because that review covered articles published before 2006, we updated the search with articles published from 2006 (to allow late inclusion in databases) to September 2011.

Selection criteria for studies

We searched for observational and randomized studies, published in English, French, Spanish or Portuguese that evaluated the associations between breastfeeding and the following outcomes: blood pressure; total cholesterol; overweight/obesity; type-2 diabetes; and performance in intelligence tests. Because we were assessing the long-term effects of breastfeeding, studies in infants were excluded.

Studies that did not use an internal comparison group were excluded. We did not apply any restrictions on the type of categorization of breastfeeding (never versus breastfed, breastfed for more or less than a given number of months, exclusively breastfed for more or less than a given number of months). Instead, as discussed in the previous section, the type of categorization of breastfeeding was considered as possible source of heterogeneity among the studies.

Type of outcome measures

In the present systematic review and meta-analysis, we searched for manuscripts that have assessed the following outcomes:

- Blood pressure: mean difference (in mmHg) in systolic and diastolic blood pressure;
- Cholesterol: mean difference (in mg/dL) in total cholesterol;
- Overweight and obesity: odds ratio comparing breastfed and non-breastfed subjects;
- Type-2 diabetes: odds ratio comparing breastfed and non-breastfed subjects;
- Performance in intelligence tests: mean difference in performance in intelligence (developmental) tests.

Search strategy

In order to minimize the likelihood of selection bias, we tried to identify as many relevant studies as possible. We carried two independent literature searches, using the terms described below. Initially, we searched Medline (2006 to September 2011) using the following terms for breastfeeding: breastfeeding; breast feeding; breastfed; breastfeed; bottle feeding; bottle fed; bottle feed; infant feeding; human milk; formula milk; formula feed; formula fed; weaning.

Every breastfeeding term was combined with each of the following terms for the outcomes:

- Cholesterol: cholesterol; LDL; HDL; triglycerides; or blood lipids;
- Type-2 diabetes: diabetes; glucose; or glycemia;
- Intellectual performance: schooling; development; or intelligence;
- Blood pressure: blood pressure; hypertension; systolic blood pressure; or diastolic blood pressure;
- Overweight or obesity: overweight; obesity; body mass index; growth; weight; height; child growth.

The titles and abstracts of studies initially selected were scanned to exclude those that were obviously irrelevant. The full text of the remaining studies was retrieved and relevant articles were identified. In addition to the electronic search, reference lists of the articles identified was searched, and we perused the Web of Science Citation Index for manuscripts citing the identified articles. Attempts were made to contact the authors of all studies that did not provide sufficient data to estimate the pooled effect. We also contacted the authors to clarify any queries on the study methodology or result.

Review methods

In this section we cover methodological aspects related to assessment of study quality, data abstraction and analysis.

Assessment of study characteristics

The following study characteristics were abstracted. These were used for two purposes, namely assessing study quality and examining reasons for heterogeneity among studies. Some studies did not provide information on all characteristics.

TABLE 4.1

Characteristics abstracted from each study

Characteristic	Categorization
Sample size	Continuous
Follow-up rates (if applicable)	Continuous
Type of study	Randomized trial Birth cohort Other
Length of recall of breastfeeding duration	< or >= 3 years
Categorization of breastfeeding	Never versus ever < or >= than X months (any breastfeeding) < or >= than X months (exclusive breastfeeding)
Source of breastfeeding information	Records Interview with the subject Interview with the mother
Control for confounding	None Socioeconomic and demographic variables Socioeconomic, demographic and maternal variables (anthropometry, intellectual stimulation, intelligence, etc, depending on the outcome under study)
Control for potential mediating variables	Yes No
Type of study population	Low income Middle income High income
Year of birth of subjects	Continuous
Age at outcome assessment	Continuous

Data abstraction

Data on the above characteristics were extracted from each study using a standardized protocol. This information was collected by two independent reviewers, with disagreements being resolved by consensus rating.

Data analysis

Pooled effect estimates

In the meta-analyses, effect measures were presented as weighted mean differences for continuous outcomes and pooled odds ratio for dichotomous outcomes. Definition of exposure to breastfeeding followed the classification used in each study. A negative mean difference denoted that breastfed subjects presented a lower value of continuous variables, whereas for dichotomous variables an odds ratio < 1 indicated that breastfed subjects presented a lower odds of the outcome.

Fixed or random-effects model

To pool the studies estimates, we used a fixed and a random-effects model. Under the fixed-effect model, we assume that there is one true effect size and, therefore, the difference among studies results is due to random variation. In the fixed-effect model, studies are weighted by their precision (inverse of the standard error) (1). On the other hand, under the random-effects model we assume that the true effects also vary, and thus the pooled effect needs to take into consideration an additional source of variation. In the random-effect model, studies are weighted by their precision plus the estimate of the between studies variance (heterogeneity) (2). By incorporating a second source of variability (variance between studies) in the estimate of the variance, the confidence interval in the random-effect model is wider than that for the fixed-effect model. Because the between studies variance is the same for every study, the random-effect model gives greater weight to smaller studies as compared to the fixed-effect model.

In the present meta-analyses, heterogeneity among studies was assessed with the Q-test and I-square; if either method suggests that between-studies variability was higher than that expected by chance, a random-effects model was used (2). Otherwise, a fixed-effect model is recommended. In this series of meta-analyses, heterogeneity was evident for all outcomes, and thus random-effects models were used throughout.

Publication bias

Studies reporting statistically significant associations are more likely to be published, and to be cited by others articles. Therefore, these results tend to be included in systematic-review, whereas small studies with negative findings are less often published. Publication bias is more likely to affect small studies because the great amount of resources (time and money) spent in larger studies makes them more likely to be published, regardless of their results (1). In meta-analysis, publication bias is a type of selection bias.

In the present meta-analyses, funnel plots and Egger's test were employed to assess the presence of publication bias (3). Furthermore, the analysis was stratified according to study size, in order to assess the impact of publication bias on the pooled estimate.

Assessing heterogeneity

The last phase of the analyses relied on meta-regression to assess the contribution of study characteristics to between-study variability (4). In this approach, if the data are homogenous or if the heterogeneity is fully explained by the covariates, the random-effects model is reduced to a fixed-effect model. This analysis was performed using the METAREG command within STATA. Each of the items listed in **Table 4.1** were included as covariates in the meta-regression, one at a time, rather than using an overall score. This approach allows the identification of aspects of study design that were responsible for heterogeneity between studies (5).

References

- 1 Greenland S. Quantitative methods in the review of epidemiologic literature. *Epidemiol Rev* 1987; 9: 1–30.
- 2 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; 7: 177–88.
- 3 Egger M, Smith GD. Bias in location and selection of studies. *BMJ* 1998; 316: 61–6.
- 4 Berkey CS, Hoaglin DC, Mosteller F, et al. A random-effects regression model for meta-analysis. *Stat Med* 1995; 14: 395–411.
- 5 Greenland S. Invited commentary: a critical look at some popular meta-analytic methods. *Am J Epidemiol* 1994; 140: 290–6.

Overweight and obesity

Overweight/obesity increases the risk of several non-communicable diseases, including diabetes, cancer and cardiovascular disease. It has been suggested that breastfeeding may prevent the development of overweight/obesity, not only in early life but also on the long-term (1).

Biological plausibility

Several mechanisms have been proposed for a protective effect of breastfeeding against obesity. Protein intake as well as energy metabolism are lower among breastfed subjects (2), and it has been suggested that higher protein intake in infancy is associated to the development of later obesity (3).

Another possibility is that breastfed and formula-fed infants have different hormonal responses to feeding, with formula leading to a greater insulin response resulting in fat deposition and increased number of adipocytes (4).

Finally, it has been proposed that differences in dietary preferences could explain such association. Scholtens et al (5) reported that at 7 years of age, Dutch children who had been breastfed for more than 16 weeks had a higher intake of fruit and vegetables than those who were never breastfed. The latter were also less likely to consume white bread, soft drinks, chocolate bars and fried snacks. However, the same study showed that adjustment for diet at 7 years did not change the magnitude of the association between breastfeeding and obesity, thus suggesting that healthier diet among breastfed subjects is not an important pathway.

Specific methodological issues

Methodological issues affecting studies of the long-term consequences of breastfeeding were addressed in [Chapter 2](#). As will be discussed in greater detail in [Chapter 9](#) (breastfeeding and intelligence), breastfeeding mothers are likely to be more health-conscious, and, therefore, to promote healthy habits, which are likely to prevent overweight and obesity later in childhood.

In addition, two methodological issues should be taken into account when studying overweight/obesity as the outcome – use of continuous BMI or of prevalence of overweight or obesity as the outcome, and adjustment for maternal BMI.

Breastfeeding may reduce the variance of weight distribution, decreasing the prevalence of both overweight and underweight in later life (6,7). As a consequence, the mean body mass index may not be affected by breastfeeding, in spite of a reduction in the prevalence of overweight/obesity. Studies that only reported an effect on mean BMI were not included in the present meta-analyses, because it would not be possible to combine their results with those reported here. Nevertheless, a previous meta-analysis by Owen et al (8) found a very small effect of breastfeeding on mean BMI (-0.04 kg/m^2) in such studies.

Maternal pregestational overweight/obesity is negatively associated with breastfeeding incidence and duration (9–13) and positively associated with offspring weight (14). For this reason, confounding by maternal anthropometry may overestimate the benefit of breastfeeding.

Overview of existing meta-analysis

In 2007, we carried out a systematic review and meta-analysis that assessed the long-term consequences of breastfeeding duration, in which overweight/obesity was one of the outcomes assessed. In that review, subjects who had been breastfed were less likely to be overweight and/or obese [pooled odds ratio: 0.78 (95% confidence interval: 0.72; 0.84)]. This association was not modified by the presence of control for confounding, year at birth, and age at assessment of body weight. Publication bias was observed, but the pooled effect among large studies (≥ 1500 participants) was similar to that observed in the overall analysis, suggesting that the publication bias did not overestimate of the mean effect

Update of the 2007 meta-analysis

In the present update, the search strategy was the same as in the systematic review published in 2007. We identified 33 new manuscripts that reported the effect of breastfeeding on the prevalence of overweight and/or obesity. Below, we will describe the findings from those new studies whose sample size was > 2000 participants. Smaller studies were also included in the meta-analyses but are not described individually.

Al-Qaoud et al (15) evaluated Kuwaiti pre-school children aged 3–6 years, and reported that those who had been breastfed presented higher odds of obesity [1.29 (95% confidence interval: 0.88; 1.90)], but the confidence interval included the unity.

Neutzling et al (16) evaluated subjects who have been followed since birth in 1993, in Pelotas, a southern Brazilian city. At 11 years, obesity was more frequent among those children who had been breastfed [odds ratio: 1.34 (95% confidence interval: 0.45; 3.91)]. As in the Kuwaiti study, the confidence interval included the unity.

Procter et al (17) linked data on Kansas families from the Pediatric Nutrition Surveillance System and Pregnancy Nutrition Surveillance System, from 1998 to 2002. Breastfeeding for ≥ 9 weeks was not related to overweight/obesity at 4 years [odds ratio: 0.95 (95% confidence interval: 0.75; 1.22)].

Huus et al (18) analyzed data from a birth cohort study from Southeast Sweden. At 5 years of age, children who had been exclusively breastfed for ≥ 4 months were somewhat less likely to be considered as obese [odds ratio: 0.82 (95% confidence interval: 0.55; 1.23)] than those who were exclusively breastfed for < 4 months. Again, the confidence interval included the unity.

Scholtens et al (19) evaluated 7 year-old children who had been enrolled in a birth cohort study primarily aimed at assessing asthma and mite allergy. Children who had been breastfed for > 16 weeks were less likely to be overweight/obese [odds ratio: 0.76 (95% confidence interval: 0.52; 1.11)], a non-significant difference.

In an English birth cohort study, Toshke et al (20) observed that the odds ratio of overweight/obesity was of 1.14 (95% confidence interval: 0.91; 1.43) among those children who had been breastfed for ≥ 6 months in relation to those who had never been breastfed.

Shields et al (21) analyzed data from a birth cohort in Brisbane, Australia, reporting that obesity at 21 years was not related to duration of breastfeeding.

Michels et al (22) reported that among women aged 37–54 years, the odds of obesity was similar among subjects who had been exclusively breastfed for > 6 months, compared to those who had never been breastfed [odds ratio: 0.94 (95% confidence interval: 0.83; 1.07)].

Fall et al (23) in the COHORTS collaboration, a consortium comprising five birth cohort studies from low and middle income countries (Brazil, Guatemala, India, Philippines and South Africa), observed that those subjects who were ever breastfed were less likely to be overweight/obese [odds ratio: 0.84 (95% confidence interval: 0.67; 1.06)], a non-significant difference.

Our meta-analyses covered 71 manuscripts that provided 75 estimates on the association between breastfeeding and overweight/obesity (**Table 5.1**). For cohort studies that provided estimates on the effects of breastfeeding for the same subjects at different age ranges, we included only the most recent follow-up in the main meta-analysis. On the other hand, in the meta-analyses stratified by age groups, we used more than one report from the same study if the results referred to different age ranges. **Figure 5.1** shows that the studies were clearly heterogeneous, and a random-effects model was thus used. Subjects who were breastfed were less likely to be considered as overweight/obese [pooled odds ratio: 0.76 (95% confidence interval: 0.71; 0.81)].

Table 5.2 shows that there was no effect modification by length of recall of breastfeeding, study setting or categorization of breastfeeding. On the other hand, the protective effect of breastfeeding was smaller in studies that assessed the prevalence of overweight/obesity in adults [pooled odds ratio: 0.89 (95% confidence interval: 0.84; 0.96)] rather than in children or adolescents. Other study characteristics also modified the association: the protective effect of breastfeeding was smaller in cohort studies, in studies whose subjects were born before 1980 and in those including adjustment for socioeconomic status, birth condition (birthweight or gestation age) and parental anthropometry. In the meta-regression analysis, three variables explained part of the variance among studies: study size (38.8% of the variance); study design (18.1% of the variance); and age at assessment of BMI (12.6% of the variance).

We further explored the influence of study design in a meta-regression analysis. When study size was not in the model, cohort studies produced odds ratios estimates that were 20% (95% confidence interval: 3%; 41%) higher than cross-sectional studies; the corresponding odds ratios were 0.83 and 0.68, indicating less protection in cohort than cross-sectional studies. After controlling for study size this ratio was reduced to 12% (95% confidence interval: -2%; 28%). Therefore, the observed effect modification by study design was partially explained by differences in sample size, with cohort studies in general being larger than case-control studies.

The presence of effect modification by study size, with small studies reporting larger effects of breastfeeding, suggests the possibility of publication bias. Indeed, the funnel plot is quite asymmetrical (**Figure 5.2**) Egger's test was statistically significant (p-value 0.003).

Our meta-analysis compared groups of children classified according to breastfeeding practices. We did not include the only general population controlled trial on this issue, in which maternity hospitals in Belarus were randomized to receive breastfeeding promotion or no intervention. It was not possible to incorporate results from this study because they compared two groups using intent-to-treat analyses, rather than comparing breastfed versus non breastfed children. This study showed no difference between the intervention and control arms at the age of 6.5 years in terms of BMI or adiposity measures (24).

Conclusion

In this updated meta-analysis, we observed an association between breastfeeding and lower prevalence of overweight/obesity later in life. Nevertheless, some methodological issues should be taken into consideration in the assessment of the evidence. A small-study effect (publication bias) is an important issue, tending to overestimate the benefits of breastfeeding. Among studies with ≥ 1500 participants the protective effect of breastfeeding was relatively modest [0.85 (95% confidence interval: 0.80; 0.91)].

Residual confounding is another issue that should be addressed, because most studies were carried out in high-income countries where breastfeeding tends to be more common among the better off and more educated mothers. In these societies, overweight and obesity tend to be more prevalent among the poor, and even studies that adjusted for several socioeconomic variables may still be affected by residual confounding. Attempting to elucidate this possibility, Brion et al (25) compared the effects of breastfeeding on body mass index in two settings with different socioeconomic confounding structures. In England, a developed country setting, breastfeeding was protective against overweight, but in Brazil, where breastfeeding does not show a clear social gradient, no such effect was evident. This was confirmed by the negative findings of the COHORTS collaboration from low and middle-income countries (23). Therefore, residual confounding by socioeconomic status is an issue that should be taken into consideration in the assessment of causality. By the same token, we observed that studies with tighter control of confounding (socioeconomic factors, birth weight or gestational age, and parental anthropometry) reported smaller benefits of breastfeeding.

In order to reduce the influence of publication bias and residual confounding, we estimated the pooled effect among those studies with large sample size and that controlled for confounding by socioeconomic, birth weight or gestational age, and parental anthropometry. The pooled odds ratio in the 16 studies that fulfilled both criteria was 0.88 (95% confidence interval: 0.83; 0.93), suggesting that the protective effect of breastfeeding may be overestimated by publication bias and residual confounding.

Our conclusion is that the meta-analysis of higher-quality studies suggests a small reduction, of about 10%, in the prevalence of overweight or obesity in children exposed to longer durations of breastfeeding. Nevertheless, it is not possible to completely rule out residual confounding because in most study settings breastfeeding duration was higher in families where the parents were more educated and had higher income levels.

TABLE 5.1
Breastfeeding and overweight/obesity in later life: studies included in the meta-analysis in ascending order of subjects age at which outcome was measured

Author, Year	Year of birth of subjects	Study design	Mean age at measurement	Gender	Comparison groups	Outcome	Odds ratio (95% confidence interval)
Jingxiong, 2009 (26)	2003–2004	Case-control	1 year	All	Exclusively breastfed for > 4 months vs. Exclusively breastfed for ≤ 4 months	Overweight or obesity	0.47 (0.25; 0.87)
Weyermann, 2006 (27)	2000–2001	Cohort	2 years	All	Breastfed for ≥ 6 months vs. Breastfed for < 3 months	Overweight or obesity	0.4 (0.2; 0.8)
Poulton, 2001 (28)	1972–1973	Cohort	3 years	All	Breastfed for > 6 months vs. Never breastfed	Overweight or obesity	0.86 (0.44; 1.65)
Armstrong, 2002 (29)	1995–1996	Cohort	3 years	All	Exclusively breastfed vs. Bottle fed at 6–8 months	Obesity	0.72 (0.65; 0.79)
Taveras, 2006 (30)	Not stated	Cohort	3 years	All	No infant formula feeding during the first 6 months of life vs. Not breastfed at 6 months	Overweight or obesity	0.33 (0.13; 0.84)
Dennison, 2006 (31)	1995–1999	Cross-sectional	4 years	All	Breastfed ≥ 1 month vs. Never breastfed	Obesity	0.73 (0.48; 1.13)
Strbak, 1991 (32)	Not stated	Cohort	4 years	All	Breastfed for > 1 month vs. Breastfed < 2 weeks	Obesity	0.84 (0.41; 1.72)
Hediger 2001 (33)	1983–1991	Cross-sectional	4 years	All	Ever breastfed vs. Never breastfed	Obesity	0.84 (0.62; 1.13)
Grummer-Strawn, 200 (46)	1988–1992	Cohort	4 years	All	Breastfed for ≥ 12 months vs. Never breastfed	Overweight or obesity	0.72 (0.65; 0.80)
Li, 2005 (34)	1990–1994	Cohort	4 years	All	Breastfed for ≥ 4 months vs. Never breastfed	Obesity	0.6 (0.3; 1.3)
Dubois, 2006 (35)	1998	Cohort	4 years	All	Breastfed for ≥ 3 months vs. Breastfed for < 3 months	Obesity	1.0 (0.7; 1.5)
Araujo, 2006 (36)	1993	Cohort	4 years	All	Ever breastfed vs. Never breastfed	Overweight or obesity	1.83 (0.53; 6.28)
Moschonis, 2008 (37)	1998–2001	Cross-sectional	4 years	All	Exclusive breastfeeding for ≥ 6 months vs. Never breastfed	Overweight or obesity	0.94 (0.65; 1.35)
Procter, 2008 (17)	1998	Cohort	4 years	All	Breastfed for ≥ 9 months vs. Never breastfed	Overweight or obesity	0.95 (0.75; 1.22)
Al-Qaoud, 2009 (15)	Not stated	Cross-sectional	4 years	All	Ever breastfed vs. Never breastfed	Obesity	1.29 (0.88; 1.90)
Komatsu, 2009 (38)	2002–2006	Cross-sectional	4 years	All	Exclusively breastfed for ≥ 6 months vs. Exclusively breastfed for < 6 months	Overweight or obesity	0.46 (0.15; 1.37)
Simon, 2009 (39)	1988–2003	Cross-sectional	4 years	All	Exclusively breastfed for ≥ 6 months vs. Exclusively breastfed for < 6 months	Overweight or obesity	0.57 (0.38; 0.86)
Balaban, 2010 (40)	Not stated	Case-control	4 years	All	Breastfed for ≥ 4 months vs. Breastfed for < 4 months	Overweight or obesity	0.70 (0.43; 1.16)

Author, Year	Year of birth of subjects	Study design	Mean age at measurement	Gender	Comparison groups	Outcome	Odds ratio (95% confidence interval)
Gewa, 2010 (41)	2002–2006	Cross-sectional	4 years	All	Breastfed for > 24 months vs. Breastfed for < 12 months	Overweight or obesity	0.55 (0.32; 0.94)
Twells, 2010 (42)	2001	Cross-sectional	4 years	All	Exclusively breastfed for ≥ 3 months vs. Never breastfed	Obesity	0.66 (0.45; 0.97)
O'Callaghan, 1997 (43)	1981–1984	Cohort	5 years	All	Ever breastfed vs. Never breastfed	Obesity	0.71 (0.43; 1.25)
von Kries, 1999 (44)	1991–1992	Cross-sectional	5 years	All	Ever breastfed vs. Never breastfed	Obesity	0.75 (0.57; 0.98)
He, 2000 (45)	1989–1993	Case control	5 years	All	Breastfed vs. Never breastfed	Obesity	1.18 (0.91; 1.54)
Scaglioni, 2000 (46)	1991	Cohort	5 years	All	Ever breastfed vs. Never breastfed	Overweight or obesity	0.66 (0.25; 1.78)
Burdette, 2006 (47)	Not stated	Cohort	5 years	All	Ever breastfed vs. Never breastfed	Overweight or obesity	0.79 (0.44; 1.46)
Schaefer-Graf, 2006 (48)	1995–2000	Cross-sectional	5 years	All	Breastfed for > 3 months vs. Breastfed for ≤ 3 months	Overweight or obesity	0.55 (0.33; 0.91)
Huus, 2008 (18)	1997–1999	Cohort	5 years	All	Exclusively breastfed for ≥ 4 months vs. Exclusively breastfed for < 4 months	Obesity	0.82 (0.55; 1.23)
Wadsworth, 1999 (49)	1946	Cohort	6 years	All	Ever breastfed vs. Never breastfed	Obesity	0.83 (0.65; 1.04)
Thorsdottir, 2003 (50)	Not stated	Cohort	6 years	Male	Breastfed for ≥ 6 months vs. Breastfed for < 6 months	Overweight or obesity	0.33 (0.13; 0.83)
Li, 2003 (51)	1983–1987	Cross-sectional	6 years	All	Breastfed for > 9 months vs. Breastfed for < 1 week	Obesity	0.61 (0.28; 1.32)
Bergmann, 2003 (52)	1990	Cohort	6 years	All	Breastfed for ≥ 3 months vs. Breastfed for < 3 months	Obesity	0.46 (0.23; 0.92)
Richter, 1981 (53)	1957–1959	Cohort	7 years	All	Ever breastfed vs. Never breastfed	Obesity	0.73 (0.51; 1.07)
Scholtens, 2007 (19)	1996–1997	Cohort	7 years	All	Breastfed for > 16 weeks vs. Breastfed for ≤ 16 weeks	Overweight or obesity	0.76 (0.52; 1.11)
Buyken, 2008 (54)	1984–1999	Cohort	7 years	M F	Breastfed for > 4 months vs. Never breastfed Breastfed for > 4 months vs. Never breastfed	Overweight or obesity Overweight or obesity	0.70 (0.27; 1.74) 1.11 (0.50; 2.44)
Kwok, 2010 (55)	1997	Cohort	7 years	All	Exclusively breastfed for ≥ 3 months vs. Never breastfed	Overweight of obesity	1.09 (0.83; 1.43)
Eid, 1970 (56)	1961	Cohort	8 years	All	Breastfed vs. Bottle fed	Obesity	0.42 (0.11; 1.61)
Maffeis, 1994 (57)	Not stated	Cross-sectional	8 years	All	Not stated	Obesity	0.91 (0.69; 1.21)
Liese, 2001 (58)	1985–1987	Cross-sectional	9 years	All	Ever breastfed vs. Never breastfed	Overweight or obesity	0.66 (0.52; 0.87)
Frye, 2003 (59)	1978–1994	Cross-sectional	9 years	All	Exclusively breastfed for > 12 weeks vs. Never breastfed	Obesity	0.60 (0.40; 0.90)

Author, Year	Year of birth of subjects	Study design	Mean age at measurement	Gender	Comparison groups	Outcome	Odds ratio (95% confidence interval)
Mai, 2007 (60)	1995	Case-control	9 years	All	Exclusively breastfed for ≥ 12 weeks vs. Exclusively breastfed for < 12 weeks	Overweight or obesity	0.55 (0.39; 0.77)
Toschke, 2007 (20)	1991–1992	Cohort	9 years	All	Breastfed for ≥ 6 months vs. Never breastfed	Obesity	1.40 (0.92; 2.14)
Toschke, 2002 (61)	1977–1985	Cross-sectional	10 years	All	Ever breastfed vs. Never breastfed	Obesity	0.80 (0.66; 0.96)
Sung, 2003 (62)	1990	Cross-sectional	10 years	All	Breastfed vs. Bottle fed	Overweight or obesity	0.45 (0.27; 0.78)
Siqueira, 2007 (63)	1990–1998	Cross-sectional	10 years	All	Ever breastfed vs. Never breastfed	Obesity	0.49 (0.24; 0.98)
Poulton, 2001 (28)	1972–1973	Cohort	11 years	All	Breastfed for > 6 months vs. Never breastfed	Overweight or obesity	0.36 (0.1; 1.28)
Panagiotakos, 2008 (64)	Not stated	Cross-sectional	11 years	Male	Breastfeeding for > 3 months vs. Never breastfed	Overweight or obesity	0.28 (0.09; 0.84)
				Female	Breastfeeding for > 3 months vs. Never breastfed	Overweight or obesity	0.19 (0.06; 0.65)
Neutzling, 2009 (16)	1993	Cohort	11 years	All	Ever breastfed vs. Never breastfed	Obesity	1.34 (0.45; 3.91)
Sabanayagam, 2009 (65)	Not stated	Cohort	11 years	All	Ever breastfed vs. Never breastfed	Overweight or obesity	1.14 (0.80; 1.63)
Papandreou, 2010 (66)	Not stated	Cross-sectional	11 years	All	Breastfeeding for ≥ 3 months vs. Breastfed for < 3 months	Overweight or obesity	0.21 (0.11; 0.79)
Gillman 2001 (7)	Not stated	Cross-sectional	12 years	All	Mostly or only breastfed vs. Mostly or only formula fed	Obesity	0.78 (0.66; 0.91)
Li, 2005 (34)	1982–1985	Cohort	12 years	All	Breastfed for ≥ 4 months vs. Never breastfed	Obesity	0.6 (0.3; 1.6)
Fallahzadeh, 2009 (67)	Not stated	Cross-sectional	12 years	All	Breastfeeding for ≥ 24 months vs. Breastfeeding for < 12 months	Obesity	0.53 (0.31; 1.01)
Li, 2003 (51)	1973–1982	Cross-sectional	13 years	All	Breastfed for > 9 months vs. Breastfed for < 1 week	Obesity	0.73 (0.23; 2.27)
Elliott, 1997 (68)	1977–1980	Cohort	14 years	All	Breastfed for ≥ 2 months vs. Breastfed for < 2 months	Overweight or obesity	0.51 (0.22; 1.16)
Shields, 2006 (69)	1981–1984	Cohort	14 years	All	Breastfed for ≥ 4 months vs. Never breastfed	Obesity	0.92 (0.6; 1.5)
Woo, 2008 (70)	Not stated	Cross-sectional	14 years	Black	Breastfeeding for > 4 months vs. Never breastfed	0.50 (0.20; 1.27)	0.48 (0.32; 0.74)
				White	Breastfeeding for > 4 months vs. Never breastfed		
Metzger, 2010 (71)	Not stated	Cohort	14 years	All	Ever breastfed vs. Never breastfed	Overweight or obesity	0.44 (0.03; 0.85)
Kramer, 1981 (72)	Not stated	Case-control	15 years	All	Ever breastfed vs. Never breastfed	Obesity	0.44 (0.21; 0.93)

Author, Year	Year of birth of subjects	Study design	Mean age at measurement	Gender	Comparison groups	Outcome	Odds ratio (95% confidence interval)
Kramer, 1981 (72)	Not stated	Case-control	15 years	All	Ever breastfed vs. Never breastfed	Obesity	0.29 (0.11; 0.75)
Tulldahl, 1999 (73)	1979	Cohort	15 years	All	Breastfed for > 2 months vs. Breastfed for ≤ 2 months	Overweight or obesity	0.66 (0.44; 0.98)
Victoria, 2003 (74)	1982	Cohort	18 years	Male	Breastfed for ≥ 1 month vs. Breastfed for < 1 month	Obesity	0.73 (0.50; 1.07)
Poulton, 2001 (28)	1972–1973	Cohort	21 years	All	Breastfed for > 6 months vs. Never breastfed	Overweight or obesity	0.79 (0.46; 1.34)
Shields, 2010 (21)	1981–1983	Cohort	21 years	All	Breastfed for > 4 months vs. Never breastfed	Obesity	0.93 (0.63; 1.39)
Fall, 2011 (23)	1969–1990	Cohort	25 years	All	Ever breastfed vs. Never breastfed	Overweight or obesity	0.84 (0.67; 1.06)
Parsons, 2003 (75)	1958	Cohort	33 years	Males	Breastfed for > 1 month vs. Never breastfed	Obesity	0.93 (0.74; 1.17)
Kvaavik, 2005 (76)	1966	Cohort	33 years	Females	Breastfed for > 1 month vs. Never breastfed	Obesity	0.84 (0.67; 1.05)
Parikh, 2009 (77)	Not stated	Cohort	41 years	All	Breastfed for ≥ 4 month vs. Never breastfed	Obesity	0.34 (0.12; 1.01)
Michels, 2007 (22)	1947–1964	Cohort	45 years	Female	Ever breastfed vs. Never breastfed	Obesity	0.75 (0.47; 1.21)
Rudnicka, 2007 (78)	1958	Cohort	45 years	All	Exclusively breastfed for > 6 months vs. Never breastfed	Obesity	0.94 (0.83; 1.07)
Eriksson, 2003 (79)	1934–1944	Cohort	61 years	All	Breastfed for > 1 month vs. Never breastfed	Obesity	0.85 (0.75; 0.97)
O'Tierney, 2009 (80)	1934–1944	Cohort	62 years	All	Breastfed for > 8 months vs. Breastfed for < 2 months	Obesity	1.10 (0.88; 1.37)
					Breastfed for ≥ 8 months vs. Breastfed for < 2 months		0.75 (0.50; 1.11)

TABLE 5.2

Breastfeeding and the risk of overweight and obesity in later life: Random-effects meta-analyses of risk of overweight/obesity by subgroup

Subgroup analysis	Number of estimates	Pooled odds ratio and 95% confidence interval	P Value
By age group			
1 to 9 years	42	0.77 (0.71; 0.83)	< 0.001
10 to 19 years	22	0.62 (0.53; 0.73)	< 0.001
≥ 20 years	11	0.89 (0.84; 0.96)	0.001
By study size			
< 500 participants	17	0.51 (0.42; 0.61)	< 0.001
500–1499 participants	24	0.65 (0.56; 0.76)	< 0.001
≥ 1500 participants	27	0.85 (0.80; 0.91)	< 0.001
By year at birth			
Before 1980	14	0.87 (0.80; 0.94)	< 0.001
After 1980	35	0.76 (0.70; 0.83)	< 0.001
By study design			
Cohort	38	0.83 (0.77; 0.89)	< 0.001
Case-control	6	0.60 (0.39; 0.92)	0.02
Cross-sectional	24	0.68 (0.60; 0.76)	< 0.001
By length of recall of breastfeeding			
< 3 years	29	0.80 (0.73; 0.88)	< 0.001
≥ 3 years	39	0.73 (0.67; 0.80)	< 0.001
By control for confounding			
None	17	0.66 (0.54; 0.82)	< 0.001
Adjusted for socioeconomic status	5	0.63 (0.52; 0.76)	< 0.001
Also adjusted for birth condition	12	0.77 (0.69; 0.86)	< 0.001
Also adjusted for parental anthropometry	34	0.81 (0.75; 0.88)	< 0.001
By setting			
High income country	48	0.76 (0.71; 0.81)	< 0.001
Middle /low income country	20	0.75 (0.64; 0.89)	0.001
By categorization of breastfeeding			
Ever breastfed	20	0.79 (0.70; 0.90)	< 0.001
Breastfed for a given number of months	35	0.74 (0.67; 0.82)	< 0.001
Exclusively breastfed for a given number of months	12	0.73 (0.62; 0.85)	< 0.001
Total	75	0.76 (0.71; 0.81)	

Note: The total number of studies does not add to 75, due to exclusion of studies with repeated report on the effect of breastfeeding (7 studies), one study with missing information on categorization of breastfeeding and 19 studies with missing information on year of birth of subjects.

FIGURE 5.1

Odds ratio and its 95% confidence interval of being considered as overweight/obese, comparing breastfed vs. not-breastfed subjects in different studies. Whether the estimate was for males (M), females (F) and all(A) is indicated in parenthesis.

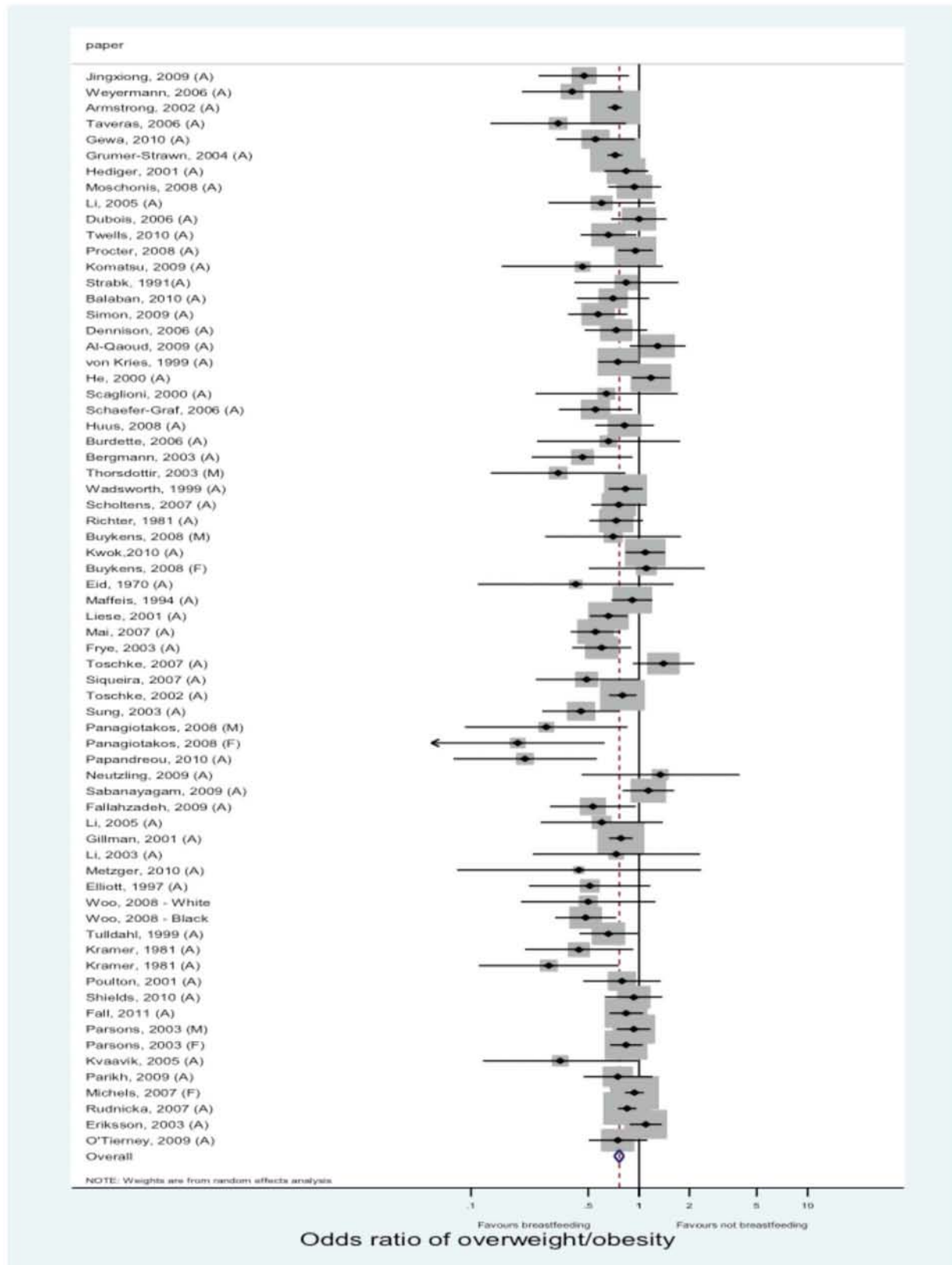
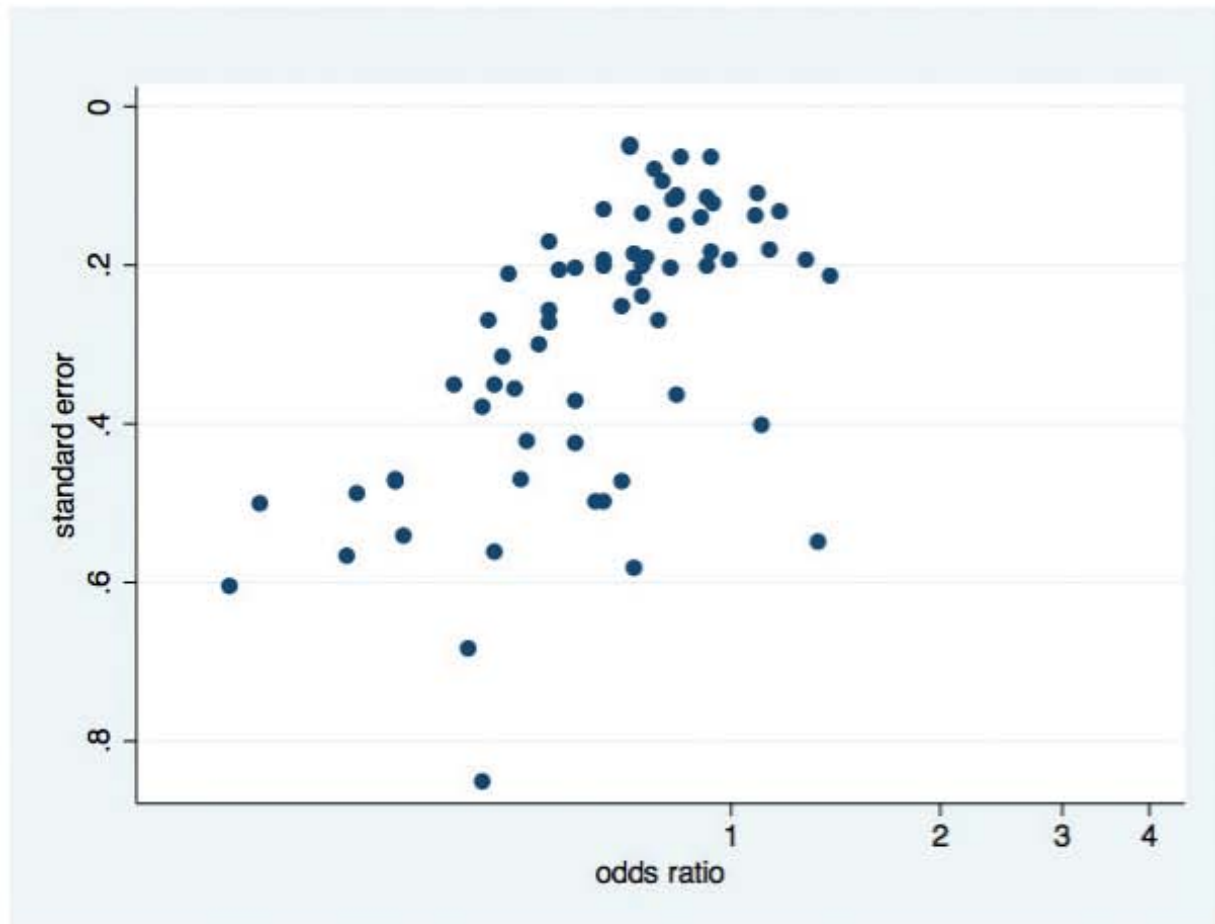


FIGURE 5.2

Funnel plot showing odds ratio for overweight/obesity by standard error of odds ratio



References

- 1 Gillman MW, Rifas-Shiman SL, Camargo CA, Jr., et al. Risk of overweight among adolescents who were breastfed as infants. *JAMA* 2001; 285: 2461–7.
- 2 Whitehead RG. For how long is exclusive breast-feeding adequate to satisfy the dietary energy needs of the average young baby? *Pediatr Res* 1995; 37: 239–43.
- 3 Rolland-Cachera MF, Deheeger M, Akrou M, et al. Influence of macronutrients on adiposity development: a follow up study of nutrition and growth from 10 months to 8 years of age. *Int J Obes Relat Metab Disord* 1995; 19: 573–8.
- 4 Lucas A, Sarson DL, Blackburn AM, et al. Breast vs bottle: endocrine responses are different with formula feeding. *Lancet* 1980; 1: 1267–9.
- 5 Scholtens S, Brunekreef B, Smit HA, et al. Do differences in childhood diet explain the reduced overweight risk in breastfed children? *Obesity* (Silver Spring) 2008; 16: 2498–503.
- 6 Grummer-Strawn LM, Mei Z. Does breastfeeding protect against pediatric overweight? Analysis of longitudinal data from the Centers for Disease Control and Prevention Pediatric Nutrition Surveillance System. *Pediatrics* 2004; 113: e81–6.
- 7 Beyerlein A, Toschke AM, von Kries R. Breastfeeding and childhood obesity: shift of the entire BMI distribution or only the upper parts? *Obesity* (Silver Spring) 2008; 16: 2730–3.
- 8 Owen CG, Martin RM, Whincup PH, et al. The effect of breastfeeding on mean body mass index throughout life: a quantitative review of published and unpublished observational evidence. *Am J Clin Nutr* 2005; 82: 1298–307.
- 9 Oddy WH, Li J, Landsborough L, et al. The association of maternal overweight and obesity with breastfeeding duration. *J Pediatr* 2006; 149: 185–91.
- 10 Li R, Jewell S, Grummer-Strawn L. Maternal obesity and breast-feeding practices. *Am J Clin Nutr* 2003; 77: 931–6.
- 11 Donath SM, Amir LH. Does maternal obesity adversely affect breastfeeding initiation and duration? *J Paediatr Child Health* 2000; 36: 482–6.
- 12 Chapman DJ, Perez-Escamilla R. Identification of risk factors for delayed onset of lactation. *J Am Diet Assoc* 1999; 99: 450–4; quiz 5–6.
- 13 Baker JL, Michaelsen KF, Sorensen TI, et al. High prepregnant body mass index is associated with early termination of full and any breastfeeding in Danish women. *Am J Clin Nutr* 2007; 86: 404–11.
- 14 Drake AJ, Reynolds RM. Impact of maternal obesity on offspring obesity and cardiometabolic disease risk. *Reproduction* 2009; 140: 387–98.
- 15 Al-Qaoud N, Prakash P. 'Can breastfeeding and its duration determine the overweight status of Kuwaiti children at the age of 3–6 years?'. *Eur J Clin Nutr* 2009; 63: 1041–3.
- 16 Neutzling MB, Hallal PR, Araujo CL, et al. Infant feeding and obesity at 11 years: prospective birth cohort study. *Int J Pediatr Obes* 2009; 4: 143–9.
- 17 Procter SB, Holcomb CA. Breastfeeding duration and childhood overweight among low-income children in Kansas, 1998–2002. *Am J Public Health* 2008; 98: 106–10.
- 18 Huus K, Ludvigsson JF, Enskar K, et al. Exclusive breastfeeding of Swedish children and its possible influence on the development of obesity: a prospective cohort study. *BMC Pediatr* 2008; 8: 42.
- 19 Scholtens S, Gehring U, Brunekreef B, et al. Breastfeeding, weight gain in infancy, and overweight at seven years of age: the prevention and incidence of asthma and mite allergy birth cohort study. *Am J Epidemiol* 2007; 165: 919–26.
- 20 Toschke AM, Martin RM, von Kries R, et al. Infant feeding method and obesity: body mass index and dual-energy X-ray absorptiometry measurements at 9–10 y of age from the Avon Longitudinal Study of Parents and Children (ALSPAC). *Am J Clin Nutr* 2007; 85: 1578–85.
- 21 Shields L, Mamun AA, O'Callaghan M, et al. Breastfeeding and obesity at 21 years: a cohort study. *J Clin Nurs* 19: 1612–7.
- 22 Michels KB, Willett WC, Graubard BI, et al. A longitudinal study of infant feeding and obesity throughout life course. *Int J Obes (Lond)* 2007; 31: 1078–85.
- 23 Fall CH, Borja JB, Osmond C, et al. Infant-feeding patterns and cardiovascular risk factors in young adulthood: data from five cohorts in low- and middle-income countries. *Int J Epidemiol* 40: 47–62.
- 24 Kramer MS, Matush L, Vanilovich I, et al. A randomized breast-feeding promotion intervention did not reduce child obesity in Belarus. *J Nutr* 2009; 139: 417S–21S.

- 25 Brion MJ, Lawlor DA, Matijasevich A, et al. What are the causal effects of breastfeeding on IQ, obesity and blood pressure? Evidence from comparing high-income with middle-income cohorts. *Int J Epidemiol*; 40: 670–80.
- 26 Jingxiong J, Rosenqvist U, Huishan W, et al. Relationship of parental characteristics and feeding practices to overweight in infants and young children in Beijing, China. *Public Health Nutr* 2009; 12: 973–8.
- 27 Weyermann M, Rothenbacher D, Brenner H. Duration of breastfeeding and risk of overweight in childhood: a prospective birth cohort study from Germany. *Int J Obes (Lond)* 2006; 30: 1281–7.
- 28 Poulton R, Williams S. Breastfeeding and risk of overweight. *JAMA* 2001; 286: 1449–50.
- 29 Armstrong J, Reilly JJ. Breastfeeding and lowering the risk of childhood obesity. *Lancet* 2002; 359: 2003–4.
- 30 Taveras EM, Rifas-Shiman SL, Scanlon KS, et al. To what extent is the protective effect of breastfeeding on future overweight explained by decreased maternal feeding restriction? *Pediatrics* 2006; 118: 2341–8.
- 31 Dennison BA, Edmunds LS, Stratton HH, et al. Rapid infant weight gain predicts childhood overweight. *Obesity* (Silver Spring) 2006; 14: 491–9.
- 32 Strbak V, Skultetyova M, Hromadova M, et al. Late effects of breast-feeding and early weaning: seven-year prospective study in children. *Endocr Regul* 1991; 25: 53–7.
- 33 Hediger ML, Overpeck MD, Kuczmarski RJ, et al. Association between infant breastfeeding and overweight in young children. *JAMA* 2001; 285: 2453–60.
- 34 Li C, Kaur H, Choi WS, et al. Additive interactions of maternal prepregnancy BMI and breast-feeding on childhood overweight. *Obes Res* 2005; 13: 362–71.
- 35 Dubois L, Girard M. Early determinants of overweight at 4.5 years in a population-based longitudinal study. *Int J Obes (Lond)* 2006; 30: 610–7.
- 36 Araújo CL, Victora CG, Hallal PC, et al. Breastfeeding and overweight in childhood: evidence from the Pelotas 1993 birth cohort study. *International journal of obesity* (2005) 2006; 30: 500–6.
- 37 Moschonis G, Grammatikaki E, Manios Y. Perinatal predictors of overweight at infancy and preschool childhood: the GENESIS study. *Int J Obes (Lond)* 2008; 32: 39–47.
- 38 Komatsu H, Yorifuji T, Iwase T, et al. Impact of breastfeeding on body weight of preschool children in a rural area of Japan: population-based cross-sectional study. *Acta Med Okayama* 2009; 63: 49–55.
- 39 Simon VG, Souza JM, Souza SB. Breastfeeding, complementary feeding, overweight and obesity in pre-school children. *Rev Saude Publica* 2009; 43: 60–9.
- 40 Balaban G, Motta ME, Silva GA. Early weaning and other potential risk factors for overweight among preschool children. *Clinics* (Sao Paulo); 65: 181–7.
- 41 Gewa CA. Childhood overweight and obesity among Kenyan pre-school children: association with maternal and early child nutritional factors. *Public Health Nutr*; 13: 496–503.
- 42 Twells L, Newhook LA. Can exclusive breastfeeding reduce the likelihood of childhood obesity in some regions of Canada? *Can J Public Health*; 101: 36–9.
- 43 O'Callaghan MJ, Williams GM, Andersen MJ, et al. Prediction of obesity in children at 5 years: a cohort study. *J Paediatr Child Health* 1997; 33: 311–6.
- 44 von Kries R, Koletzko B, Sauerwald T, et al. Breast feeding and obesity: cross sectional study. *BMJ* 1999; 319: 147–50.
- 45 He Q, Ding ZY, Fong DY, et al. Risk factors of obesity in preschool children in China: a population-based case-control study. *Int J Obes Relat Metab Disord* 2000; 24: 1528–36.
- 46 Scaglioni S, Agostoni C, Notaris RD, et al. Early macronutrient intake and overweight at five years of age. *Int J Obes Relat Metab Disord* 2000; 24: 777–81.
- 47 Burdette HL, Whitaker RC, Hall WC, et al. Breastfeeding, introduction of complementary foods, and adiposity at 5 y of age. *Am J Clin Nutr* 2006; 83: 550–8.
- 48 Schaefer-Graf UM, Hartmann R, Pawliczak J, et al. Association of breast-feeding and early childhood overweight in children from mothers with gestational diabetes mellitus. *Diabetes Care* 2006; 29: 1105–7.
- 49 Wadsworth M, Marshall S, Hardy R, et al. Breast feeding and obesity. Relation may be accounted for by social factors. *BMJ* 1999; 319: 1576.
- 50 Thorsdottir I, Gunnarsdottir I, Palsson GI. Association of birth weight and breast-feeding with coronary heart disease risk factors at the age of 6 years. *Nutr Metab Cardiovasc Dis* 2003; 13: 267–72.
- 51 Li L, Parsons TJ, Power C. Breast feeding and obesity in childhood: cross sectional study. *BMJ* 2003; 327: 904–5.

- 52 Bergmann KE, Bergmann RL, Von Kries R, et al. Early determinants of childhood overweight and adiposity in a birth cohort study: role of breast-feeding. *Int J Obes Relat Metab Disord* 2003; 27: 162–72.
- 53 Richter J. [Influence of duration of breast-feeding on body-weight-development]. *Arztl Jugendkd* 1981; 72: 166–9.
- 54 Buyken AE, Karaolis-Danckert N, Remer T, et al. Effects of breastfeeding on trajectories of body fat and BMI throughout childhood. *Obesity* (Silver Spring) 2008; 16: 389–95.
- 55 Kwok MK, Schooling CM, Lam TH, et al. Does breastfeeding protect against childhood overweight? Hong Kong's 'Children of 1997' birth cohort. *Int J Epidemiol*; 39: 297–305.
- 56 Eid EE. Follow-up study of physical growth of children who had excessive weight gain in first six months of life. *Br Med J* 1970; 2: 74–6.
- 57 Maffeis C, Micciolo R, Must A, et al. Parental and perinatal factors associated with childhood obesity in north-east Italy. *Int J Obes Relat Metab Disord* 1994; 18: 301–5.
- 58 Liese AD, Hirsch T, von Mutius E, et al. Inverse association of overweight and breast feeding in 9 to 10-y-old children in Germany. *Int J Obes Relat Metab Disord* 2001; 25: 1644–50.
- 59 Frye C, Heinrich J. Trends and predictors of overweight and obesity in East German children. *Int J Obes Relat Metab Disord* 2003; 27: 963–9.
- 60 Mai XM, Becker AB, Sellers EA, et al. The relationship of breast-feeding, overweight, and asthma in preadolescents. *J Allergy Clin Immunol* 2007; 120: 551–6.
- 61 Toschke AM, Vignerova J, Lhotska L, et al. Overweight and obesity in 6- to 14-year-old Czech children in 1991: protective effect of breast-feeding. *J Pediatr* 2002; 141: 764–9.
- 62 Sung RY, Tong PC, Yu CW, et al. High prevalence of insulin resistance and metabolic syndrome in overweight/obese preadolescent Hong Kong Chinese children aged 9–12 years. *Diabetes Care* 2003; 26: 250–1.
- 63 Scanferla de Siqueira R, Monteiro CA. [Breastfeeding and obesity in school-age children from families of high socioeconomic status]. *Rev Saude Publica* 2007; 41: 5–12.
- 64 Panagiotakos DB, Papadimitriou A, Anthracopoulos MB, et al. Birthweight, breast-feeding, parental weight and prevalence of obesity in schoolchildren aged 10–12 years, in Greece; the Physical Activity, Nutrition and Allergies in Children Examined in Athens (PANACEA) study. *Pediatr Int* 2008; 50: 563–8.
- 65 Sabanayagam C, Shankar A, Chong YS, et al. Breast-feeding and overweight in Singapore school children. *Pediatr Int* 2009; 51: 650–6.
- 66 Papandreou D, Malindretos P, Rousso I. Risk factors for childhood obesity in a Greek paediatric population. *Public Health Nutr*; 13: 1535–9.
- 67 Fallahzadeh H, Golestan M, Rezvanian T, et al. Breast-feeding history and overweight in 11 to 13-year-old children in Iran. *World J Pediatr* 2009; 5: 36–41.
- 68 Elliott KG, Kjolhede CL, Gournis E, et al. Duration of breastfeeding associated with obesity during adolescence. *Obes Res* 1997; 5: 538–41.
- 69 Shields L, O'Callaghan M, Williams GM, et al. Breastfeeding and obesity at 14 years: a cohort study. *J Paediatr Child Health* 2006; 42: 289–96.
- 70 Woo JG, Dolan LM, Morrow AL, et al. Breastfeeding helps explain racial and socioeconomic status disparities in adolescent adiposity. *Pediatrics* 2008; 121: e458–65.
- 71 Metzger MW, McDade TW. Breastfeeding as obesity prevention in the United States: a sibling difference model. *Am J Hum Biol*; 22: 291–6.
- 72 Kramer MS. Do breast-feeding and delayed introduction of solid foods protect against subsequent obesity? *J Pediatr* 1981; 98: 883–7.
- 73 Tulldahl J, Pettersson K, Andersson SW, et al. Mode of infant feeding and achieved growth in adolescence: early feeding patterns in relation to growth and body composition in adolescence. *Obes Res* 1999; 7: 431–7.
- 74 Victora CG, Barros F, Lima RC, et al. Anthropometry and body composition of 18 year old men according to duration of breast feeding: birth cohort study from Brazil. *BMJ* 2003; 327: 901.
- 75 Parsons TJ, Power C, Manor O. Infant feeding and obesity through the lifecourse. *Arch Dis Child* 2003; 88: 793–4.
- 76 Kvaavik E, Tell GS, Klepp KI. Surveys of Norwegian youth indicated that breast feeding reduced subsequent risk of obesity. *J Clin Epidemiol* 2005; 58: 849–55.
- 77 Parikh NI, Hwang SJ, Ingelsson E, et al. Breastfeeding in infancy and adult cardiovascular disease risk factors. *Am J Med* 2009; 122: 656–63 e1.

- 78 Rudnicka AR, Owen CG, Strachan DP. The effect of breastfeeding on cardiorespiratory risk factors in adult life. *Pediatrics* 2007; 119: e1107–15.
- 79 Eriksson J, Forsen T, Osmond C, et al. *Obesity* from cradle to grave. *Int J Obes Relat Metab Disord* 2003; 27: 722–7.
- 80 O'Tierney PF, Barker DJ, Osmond C, et al. Duration of breast-feeding and adiposity in adult life. *J Nutr* 2009; 139: 422S–5S.

Blood pressure

Cardiovascular diseases is the leading cause of death worldwide, and blood pressure is positively associated with the risk of stroke and ischemic heart disease (1). It has been suggested that the development of noncommunicable diseases may be programmed by exposures in early life (2,3).

Biological plausibility

Until the 1980s the sodium content of formulas in Western countries was much lower than that of breastmilk (4). Because sodium intake in infancy is positively associated to blood pressure (5), differences in sodium content between breastmilk and formula might be a potential mechanism for the programming of later blood pressure. However, evidence on such a programming effect are scarce and inconsistent (6,7). As a consequence, there is no consensus on whether the sodium content of infant diets is associated with later blood pressure.

Long-chain polyunsaturated fatty acids (LCPUFAs), such as docosahexanoic acid (DHA) and arachidonic acid (AA), are present in breastmilk but not in most brands of formula (8). These fatty acids are an important structural component of the vascular endothelium, and evidence suggests that supplementation with LCPUFAs reduces blood pressure in adult hypertensive subjects (9). In infants, the synthesis of DHA and AA is limited, and LCPUFAs levels in nonsupplemented formula-fed infants tends to be low. Therefore, LCPUFAs levels might be a potential mechanism for the long-term effect of breastfeeding on blood pressure. However, evidence on the long-term consequences of LCPUFA supplementation in infancy on blood pressure is controversial. Forsyth et al (10) reported lower blood pressure at 6 years among formula-fed children who had been assigned to receive a formula supplemented with LCPUFAs, compared to those receiving standard formula. In contrast, de Jong et al (11) reported that mean blood pressure levels at 9 years were similar in formula-fed infants supplemented with LCPUFAs and in a control group.

It has been suggested that breastfeeding protects against obesity (12). Because blood pressure is directly correlated to body weight, one would expect breastfeeding to also protect against high blood pressure. Nevertheless, [Chapter 5](#) of the present review suggests a weak association between breastfeeding and overweight or obesity. As a consequence, it is unlikely that prevention of obesity would consist an important mechanism for the association between breastfeeding and blood pressure.

Insulin-like growth factor 1 (IGF-1) is inversely related to blood pressure in adulthood (13). It has been suggested that IGF-1 later in life is positively associated with breastfeeding (14), and consequently IGF-1 programming may constitute a pathway for the programming of later blood pressure by breastfeeding.

Summing up, although several pathways have been proposed to explain why breastfeeding might protect against high blood pressure, there is limited actual evidence on such mechanisms.

Overview of the evidence

In 2007, we carried out a systematic review and meta-analysis on the long-term consequences of breastfeeding duration, with systolic and diastolic blood pressure included among the outcomes. In that review, systolic [mean difference: -1.21 mmHg; 95% confidence interval: -1.72; -0.70] and diastolic [mean difference: -0.49 mmHg; 95% confidence interval: -0.87; -0.11] blood pressure were lower among breastfed subjects. Publication bias was evident and tended to overestimate the benefit of breastfeeding, as among those studies with ≥ 1000 participants the mean difference in systolic blood pressure was of -0.59 mmHg (95% confidence interval: -1.00; -0.19).

Update of the 2007 meta-analysis

We used the same search strategy as in the 2007 systematic review. The additional studies identified in this new search are described below.

Järvisalo et al (15) assessed endothelial function and cardiovascular risk factors among subjects aged between 24 and 39 years who have been followed for 21 years. The manuscript only provided unadjusted estimates, with systolic blood pressure being lower among females who had been breastfed, with no association among males.

Holmes et al (16) reported that breastfeeding was not related to blood pressure among subjects aged between 20 and 25 years who have been followed since adolescence.

de Jonge et al (17) observed that durations of total or exclusive breastfeeding were not associated with blood pressure at 2 years of age, even after adjustment for possible confounding variables.

Naghettini et al (18) evaluated factors related to blood pressure among children aged 3 to 10 years. In crude analysis, the authors reported that systolic blood pressure was lower among subjects who were predominantly breastfed for at least 6 months.

Fall et al (19) in the COHORTS collaboration, a consortium comprising five birth cohort studies from low and middle income countries (Brazil, Guatemala, India, Philippines and South Africa), observed that blood pressure levels were similar among ever and never breastfed subjects. Additional analyses showed a U-shaped association between breastfeeding duration and both systolic blood pressure and hypertension, with the lowest mean levels observed in subjects breastfed for 3 to 6 months.

Rudnicka et al (20) evaluated subjects who have been followed since birth in 1958, in England, Wales and Scotland. At 44–45 years of age, breastfeeding (never versus > 1 month) was not associated with blood pressure.

Brion et al (21) compared the effect of breastfeeding on blood pressure in two cohorts (Pelotas, Brazil and ALSPAC, England). Because the confounding structure differs between the two sites, this approach improves causal inference by exploring the likelihood of residual confounding. In the present meta-analysis, we incorporated only the results for the Pelotas site, as the results for ALSPAC had already been included (22). In Pelotas, where there is no confounding of breastfeeding by socioeconomic position, breastfeeding duration was not associated with blood pressure at the mean age of 11 years.

Evelein et al (23) assessed the cardiovascular effect of exclusive breastfeeding. Carotid intima-media thickness, distensibility and elastic modulus at mean age of 5 years were the main outcomes. Children who were breastfed for 3 to 6 months presented higher carotid intima-media thickness. Only unadjusted results were reported, with no association between exclusive breastfeeding and blood pressure.

Our updated meta-analysis included 36 studies that provided 37 estimates on the effect of breastfeeding on blood pressure (**Table 6.1**). **Figure 6.1** and **6.2** show the forest plot for systolic and diastolic blood pressure. Because there was marked heterogeneity among studies, a random-effect model was used to pool the estimates. Systolic [mean effect -1.02 (95% confidence interval: -1.45; -0.59)] and diastolic [mean effect -0.37 (95% confidence interval: -0.71; -0.04)] blood pressure were lower among subjects who were breastfed.

Similar to the 2007 meta-analysis, publication bias was clearly present (**Figure 6.3** and **6.4**). Furthermore, **Table 6.2** shows that sample size was inversely related to the mean difference in systolic and diastolic blood pressure. Among those studies whose sample size was ≥ 1000 participants, the mean difference between breastfed and non-breastfed subjects in systolic blood pressure was of -0.48 (95% confidence interval: -0.85; -0.12). This suggests that publication bias has contributed to overestimating the benefits of breastfeeding.

Table 6.2 also shows that the mean difference was smaller among those studies that measured blood pressure at ≥ 20 years. Year of age of the subject, length of recall of breastfeeding and study setting did not modify the effect of breastfeeding on blood pressure. On the other hand, those studies that compared ever versus never breastfed subjects, rather than those breastfed for more or less than a given cutoff in months, found slightly smaller effects on systolic blood pressure levels.

Table 6.2 also shows marked effect modification by control of confounding. For systolic blood pressure, the mean difference among those studies providing only crude estimates was -1.53 (95% confidence interval: -2.51; -0.56), whereas among those studies that adjusted for socioeconomic and demographic variables the mean difference was -0.80 (95% confidence interval: -1.38; -0.21).

We also assessed the role of the study characteristics presented in **Table 6.2** on the heterogeneity among studies. Sample size was the important characteristic, accounting for 34.9% of the overall heterogeneity.

Studies not included in the present review

Khan et al (24) assessed vascular function among subjects aged 11 to 14 years who have been followed since birth in Dundee, Scotland. In the crude analysis, systolic blood pressure [-3.1 (95% confidence interval: -6.16; -0.04)] and diastolic blood pressure [-1.1 (95% confidence interval: -3.82; 1.62)] were lower among breastfed subjects. Because Wilson et al (25) reported confounder-adjusted estimates from the same cohort at ages 6 to 9 years, we did not include the results by Khan et al (24).

We also did not include the cluster randomized trial on promotion of breastfeeding by Kramer et al (26) in Belarus, in which 16 maternity hospitals and surrounding clinics were randomly assigned to receive an intervention based in the Baby-Friendly Hospital Initiative. The prevalence of predominant and exclusive breastfeeding at 3 and 6 months was higher in the intervention group, but systolic blood pressure levels at six years of age differed by only 0.2 mmHg (standard error 1.58). For diastolic blood pressure, the difference was also 0.2 mmHg (standard error 1.02). These estimates could not be included in the meta-analyses because the analysis was based in the intention to treat analysis and both groups included breastfed and non-breastfed subjects.

Conclusion

In the assessment of the evidence on the long-term consequence of breastfeeding on blood pressure, publication bias and residual confounding are two methodological issues that must be taken into consideration. With respect to publication bias, we observed that small studies provided estimates that clearly overstated the benefits of breastfeeding. Among studies with sample sizes greater than 1000 subjects, the benefit of breastfeeding was modest.

Control of confounding by socioeconomic and demographic factors led to smaller reported effects of breastfeeding. This is consistent with the study by Brion et al (21) comparing two cohorts with different socioeconomic confounding structures, which is discussed above. In addition, the only randomized trial on this issue did not find an association.

In order to estimate the impact of publication bias and confounding, we pooled the estimates of seven studies with large sample sizes (> 1000 participants) that provided estimates adjusted for socioeconomic status and demographic variables. The mean effect was – 0.71 (95% confidence interval: -1.24; -0.19) for systolic and -0.27 (-0.64; 0.09) for diastolic pressure.

These findings are consistent with a small protective effect of breastfeeding against systolic blood pressure, but residual confounding cannot be ruled out.

TABLE 6.1
Breastfeeding and blood pressure in later life: studies included in the meta-analysis in ascending order of subjects' age at which outcome was measured

Author, year	Study design	Year of birth of subjects	Mean age at blood pressure measurement	Gender	Comparison groups	Mean difference in mmHg (Standard error of mean difference)	
						Systolic blood pressure	Diastolic blood pressure
Boulton, 1981 (27)	Cross-sectional	1976–9	1 year	All	Breastfed for ≥ 3 months vs. Formula fed from < 3 months	-4.11 (SE 3.12)	Not stated
Zeman, 1981 (28)	Cross-sectional	Not stated	1 year	All	Breastfed vs. Bottle fed	-4.91 (SE 2.10)	0.33 (SE 0.90)
Jonge, 2010 (17)	Cohort	2002–2006	2 years	All	Ever breastfed vs. Never breastfed	-0.45 (SE 1.42)	-0.85 (SE 1.47)
Baranowski, 1992 (29)	Cross-sectional	1981–1983	3 years	All	Exclusively breastfed for ≥ 3 months vs. Bottle fed	-2.17 (SE 1.61)	0.48 (SE 1.42)
Forsyth, 2003 (10)	Cohort	1992	5 years	All	Breastfed vs. Formula fed	-2.2 (SE 1.6)	-3.4 (SE 1.4)
Lawlor, 2004 (30)	Cohort	1981–1984	5 years	All	Breastfed for ≥ 6 months vs. Breastfed for < 6 months	-1.19 (SE 0.4)	Not stated
Evelein, 2011 (23)	Cohort	2001	5 years	All	Exclusively breastfed > 6 months vs. Never breastfed	-0.1 (SE 1.67)	1.7 (SE 1.39)
Whincup, 1989 (31)	Cross-sectional	1979–1983	7 years	All	Exclusively breast for first 3 months vs. Bottle fed	-0.20 (SE 0.33)	-0.26 (SE 0.23)
Williams, 1992 (32)	Cohort	1972–1973	7 years	All	Exclusively breastfed (median 28 wk) vs. Bottle fed	-0.5 (SE 0.66)	-0.7 (SE 0.61)
Naghettini, 2010 (18)	Cross-sectional	1996–2003	7 years	All	Predominantly breastfed ≥ 6 months vs. Predominantly breastfed < 6 months	-1.52 (SE 0.86)	Not stated
Wilson, 1998 (25)	Cohort	1983–1986	8 years	All	Exclusively breastfed for > 15 wk vs. Bottle fed	-2.84 (SE 1.40)	-1.79 (SE 1.23)
Martin, 2004 (22)	Cohort	1991–1992	6–9 years	All	Ever breastfed vs. Never breastfed	-0.8 (SE 0.4)	-0.6 (SE 0.3)
Lucas, 1994 (33)	Randomized controlled trial	1982–1985	8 years	All	Allocated to banked breastmilk vs. Allocated to preterm formula	0.1 (SE 1.55)	-0.7 (SE 1.32)
Rona, 1996 (34)	Cross-sectional (England 1993 sample)	1983–1985	9 years	All	Exclusively breastfed for ≥ 3 months vs. Bottle fed	1.46 (SE 0.84)	2.54 (SE 0.76)
Rona, 1996 (34)	Cross-sectional (England 1994 sample)	1984–1986	9 years	All	Exclusively breastfed for ≥ 3 months vs. Bottle fed	-2.75 (SE 0.86)	-2.34 (SE 0.75)
Rona, 1996 (34)	Cross-sectional (Scotland 1994 sample)	1983–1986	9 years	All	Exclusively breastfed for ≥ 3 months vs. Bottle fed	-1.29 (SE 1.01)	-0.97 (SE 0.84)
Esposito-Del Puente, 1994 (35)	Cross-sectional	1980–1982	10 years	All	Breastfed vs. Bottle fed	-4.13 (SE 2.86)	0.11 (SE 0.86)

Author, year	Study design	Year of birth of subjects	Mean age at blood pressure measurement	Gender	Comparison groups	Mean difference in mmHg (Standard error of mean difference)	
						Systolic blood pressure	Diastolic blood pressure
British Cohort Study, 1982 (36)	Cohort	1970	10 years	All	Breastfed for ≥ 3 months vs. Never breast fed	-0.16 (SE 0.38)	-0.21 (SE 0.33)
Brion, 2011 (21)	Cohort	1993	11 years	All	Breastfed ≥ 6 months vs. Breastfed < 1 month	-0.50 (SE 1.19)	-0.29 (SE 0.93)
Lawlor, 2005 (37)	Cross-sectional	Not stated	12 years	All	Exclusively breastfed vs. Not exclusively breastfed	-1.7 (SE 0.66)	Not stated
Singhal, 200 (16)	Randomized controlled trial	1982–1985	15 years	All	Allocated to banked breastmilk vs. Allocated to preterm formula	-2.7 (SE 1.52)	-3.2 (SE 1.31)
Owen, 2003 (38)	Cross-sectional	1982–1986	15 years	All	Exclusively breastfed for ≥ 3 months vs. Never breastfed	0.32 (SE 0.64)	-0.17 (SE 0.35)
Taittonen, 1996 (39)	Cohort	1962–1974	18 years	All	Breastfed > 3 months vs. Bottle fed	-5.48 (SE 1.14)	Not stated
Williams, 1992 (32)	Cohort	1972–1973	18 years	All	Exclusively breastfed (median 28 wk) vs. Bottle fed	-2.63 (SE 1.34)	-2.66 (SE 1.36)
Kolacek, 1993 (40)	Cohort	1968–1969	21 years	All	Exclusively breastfed for ≥ 3 months vs. Never breastfed	0.00 (SE 1.96)	0.75 (SE 1.40)
Holmes, 2010 (16)	Cohort	1964–1968	22 years	All	Ever breastfed vs. Never breastfed	0.3 (SE 1.45)	1.5 (SE 1.17)
Leeson, 2001 (41)	Cross-sectional	1969–1975	24 years	All	Breastfed vs. Bottle fed	0.00 (SE 1.55)	-1.00 (0.94)
Martin, 2003 (42)	Cohort	1972–1974	25 years	All	Lowest vs. Highest quartile of dried formula consumption at age 3 months	-3.7 (SE 1.38)	-1.4 (SE 0.97)
Fall, 2011 (19)	Cohort	1969–1990	25 years	All	Ever breastfed vs. Never breastfed	-0.71 (SE 0.57)	-0.55 (SE 0.48)
Järvisalo, 2009 (15)	Cohort	1962–1977	31 years	M	Ever breastfed vs. Never breastfed	-1 (SE 1.77)	2 (SE 1.61)
				F	Ever breastfed vs. Never breastfed	-3 (SE 1.51)	-2 (SE 1.26)
Rudnicka, 2007 (20)	Cohort	1958	45 years	All	Breastfed > 1 months vs. Never breastfed	0.3 (SE 0.44)	0.15 (SE 0.30)
Ravelli, 2000 (43)	Cohort	1943–1947	51 years	All	Exclusively breastfed vs. Bottle fed	0.2 (SE 1.65)	0.9 (SE 1.08)
Martin, 2005 (44,45)	Cohort	1920–1938	52 years	Male	Breastfed vs. Bottle fed	-0.11 (SE 1.11)	-0.21 (SE 0.74)
Wadsworth, 1987 (46)	Cohort	1946	53 years	All	Exclusively breastfed for ≥ 3 months vs. Never breastfed	-0.94 (SE 0.97)	-0.32 (SE 0.37)
Fall, 1995 (47)	Cohort	1920–1930	66 years	Female	Exclusively breastfed vs. Bottle fed	-2.64 (SE 2.88)	-0.96 (SE 1.38)
Martin, 2005 (45)	Cohort	1918–1939	73 years	All	Breastfed vs. Bottle fed	-1.62 (SE 2.57)	-0.74 (SE 1.18)

TABLE 6.2
Breastfeeding and blood pressure in later life. Random-effects meta-analysis by subgroups.

Subgroup analysis	Systolic blood pressure			Diastolic blood pressure		
	Number of studies	Mean difference (95% confidence interval)	P-value	Number of studies	Mean difference (95% confidence interval)	P-value
By age group						
1 to 9 years	16	-0.98 (-1.55; -0.41)	0.001	13	-0.45 (-1.14; 0.24)	0.20
10 to 19 years	8	-1.71 (-3.00; -0.41)	0.01	6	-0.49 (-1.16; 0.19)	0.16
≥ 20 years	13	-0.54 (-1.16; 0.09)	0.09	13	-0.18 (-0.54; 0.18)	0.34
By study size						
< 300 participants	11	-2.22 (-3.24; -1.21)	< 0.001	10	-0.67 (-1.64; 0.29)	0.17
300–999 participants	15	-1.34 (-2.32; -0.37)	0.007	13	-0.41 (-1.35; 0.53)	0.39
≥ 1000 participants	11	-0.48 (-0.85; -0.12)	0.01	9	-0.26 (-0.50; -0.01)	0.04
By year of birth of subjects						
Before 1980	15	-1.07 (-1.93; -0.21)	0.02	13	-0.20 (-0.52; 0.12)	0.21
After 1980	21	-0.95 (-1.41; -0.48)	< 0.001	18	-0.50 (-1.05; 0.05)	0.08
By length of recall of breastfeeding						
< 3 years	21	-1.10 (-1.48; -0.73)	< 0.001	18	-0.54 (-0.87; -0.21)	0.001
≥ 3 years	16	-0.90 (-1.63; -0.17)	0.02	14	-0.22 (-0.76; 0.33)	0.44
By categorization of breastfeeding						
Ever breastfed	18	-0.74 (-1.13; -0.34)	< 0.001	17	-0.46 (-0.87; -0.04)	0.03
Breastfed for a given number of months	19	-1.15 (-1.82; -0.47)	0.001	15	-0.31 (-0.82; 0.21)	0.24
By control for confounding						
None	16	-1.53 (-2.51; -0.56)	0.002	13	-0.459 (-1.193; 0.274)	0.22
Adjusted for socioeconomic status	12	-0.92 (-1.71; -0.14)	0.02	12	-0.435 (-1.042; 0.173)	0.16
Adjusted for socioeconomic and demographics variables	9	-0.80 (-1.38; -0.21)	0.008	7	-0.258 (-0.624; 0.108)	0.17
Study setting						
High-income country	32	-1.02 (-1.49; -0.54)	< 0.001	28	-0.41 (-0.79; -0.04)	0.03
Middle / Low income country	5	-1.06 (-2.00; -0.12)	0.03	4	-0.27 (-1.00; 0.46)	0.47
Total	37	-1.02 (-1.45; -0.59)		32	-0.37 (-0.71; -0.04)	

Note: The total number of studies does not add to 37 for systolic blood pressure and 32 for diastolic blood pressure, due to missing information on year of birth of subjects (1 study).

FIGURE 6.1

Mean difference in systolic blood pressure in mm Hg (and its 95% confidence interval) between breastfed and non-breastfed subjects in different studies. Whether the estimate was for males (M), females (F) and all (A) is indicated in parenthesis.

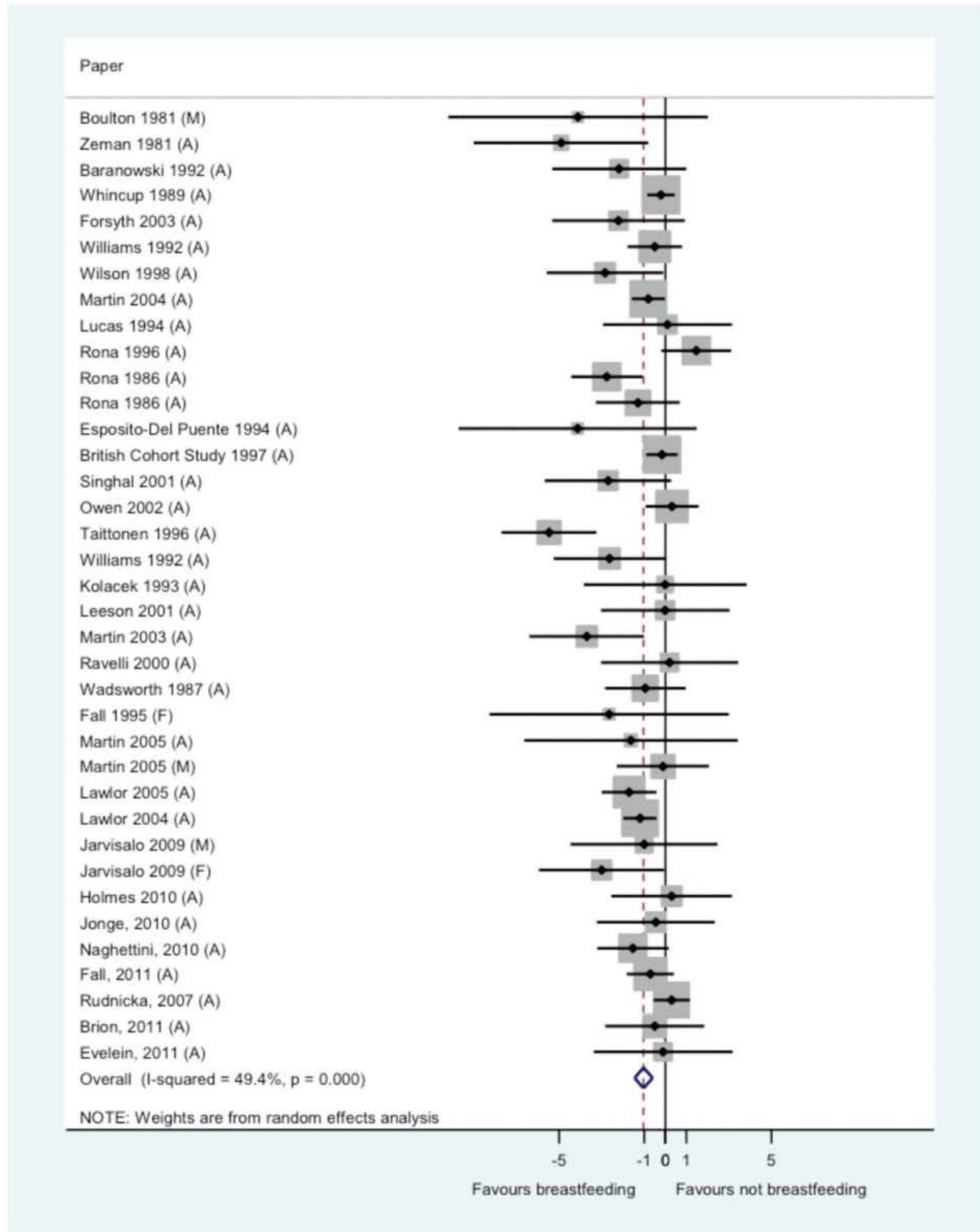


FIGURE 6.2

Mean difference in diastolic blood pressure in mm Hg (and its 95% confidence interval) between breastfed and non-breastfed subjects in different studies. Whether the estimate was for males (M), females (F) and all (A) is indicated in parenthesis.

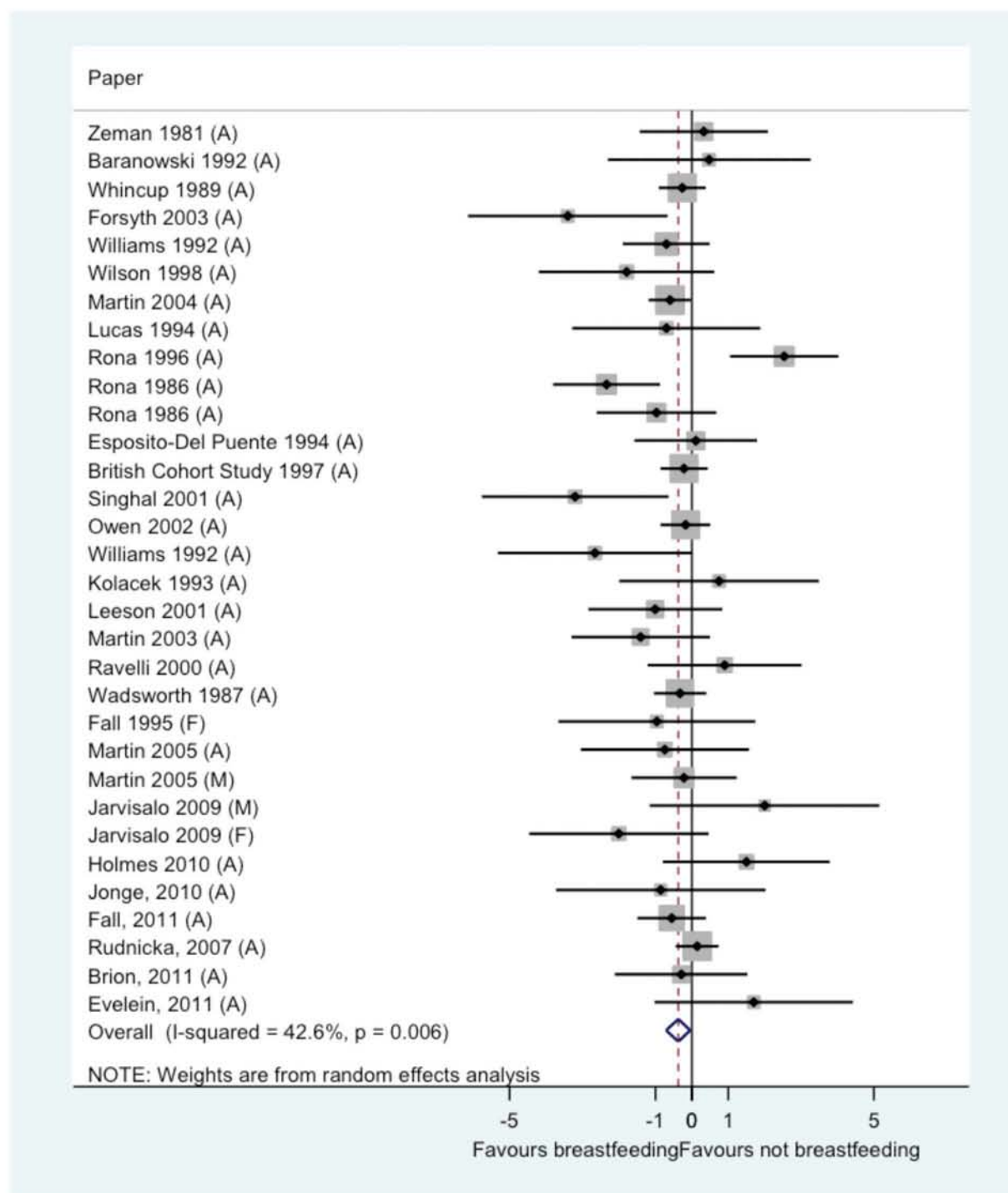


FIGURE 6.3

Funnel plot showing mean difference in systolic blood pressure (mm Hg) by standard error of mean difference

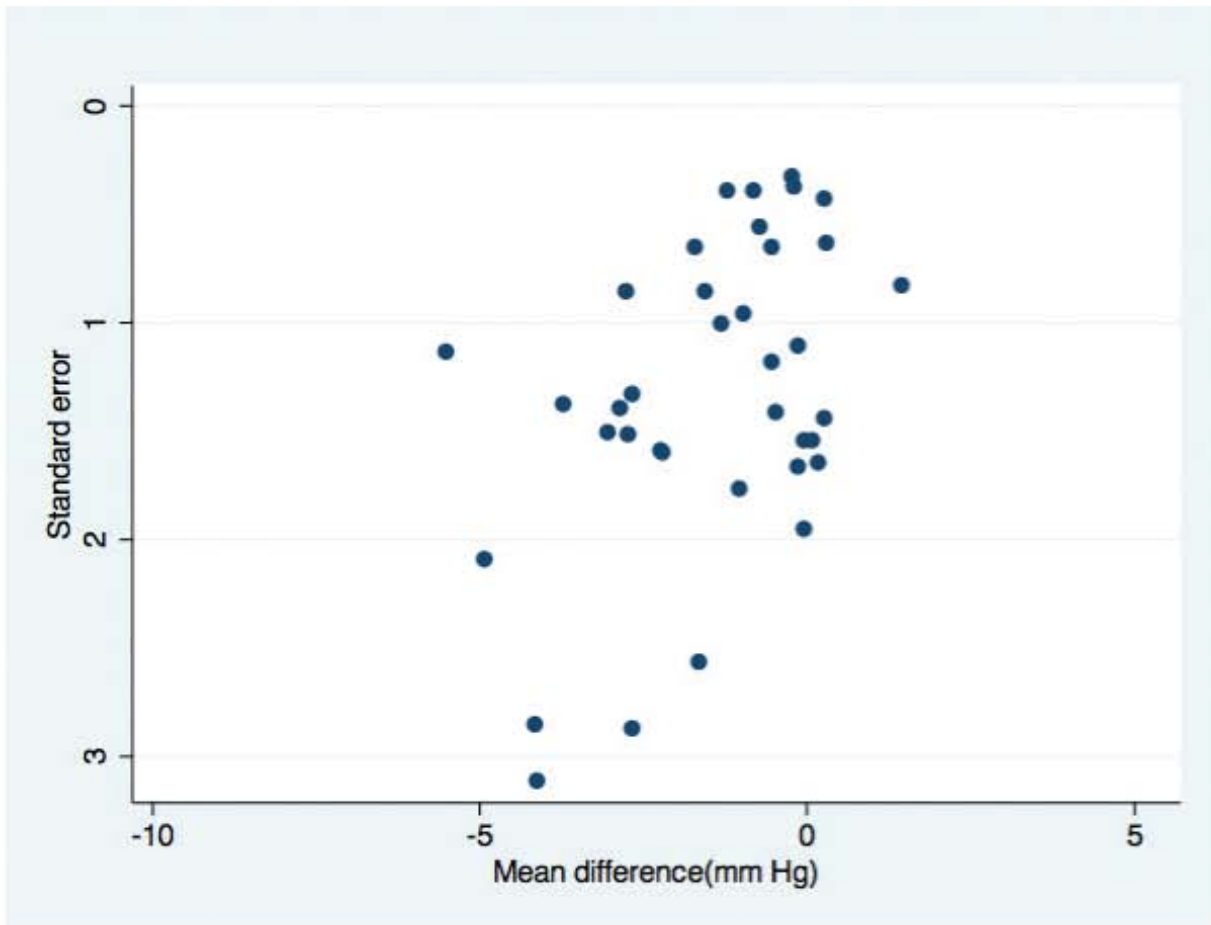
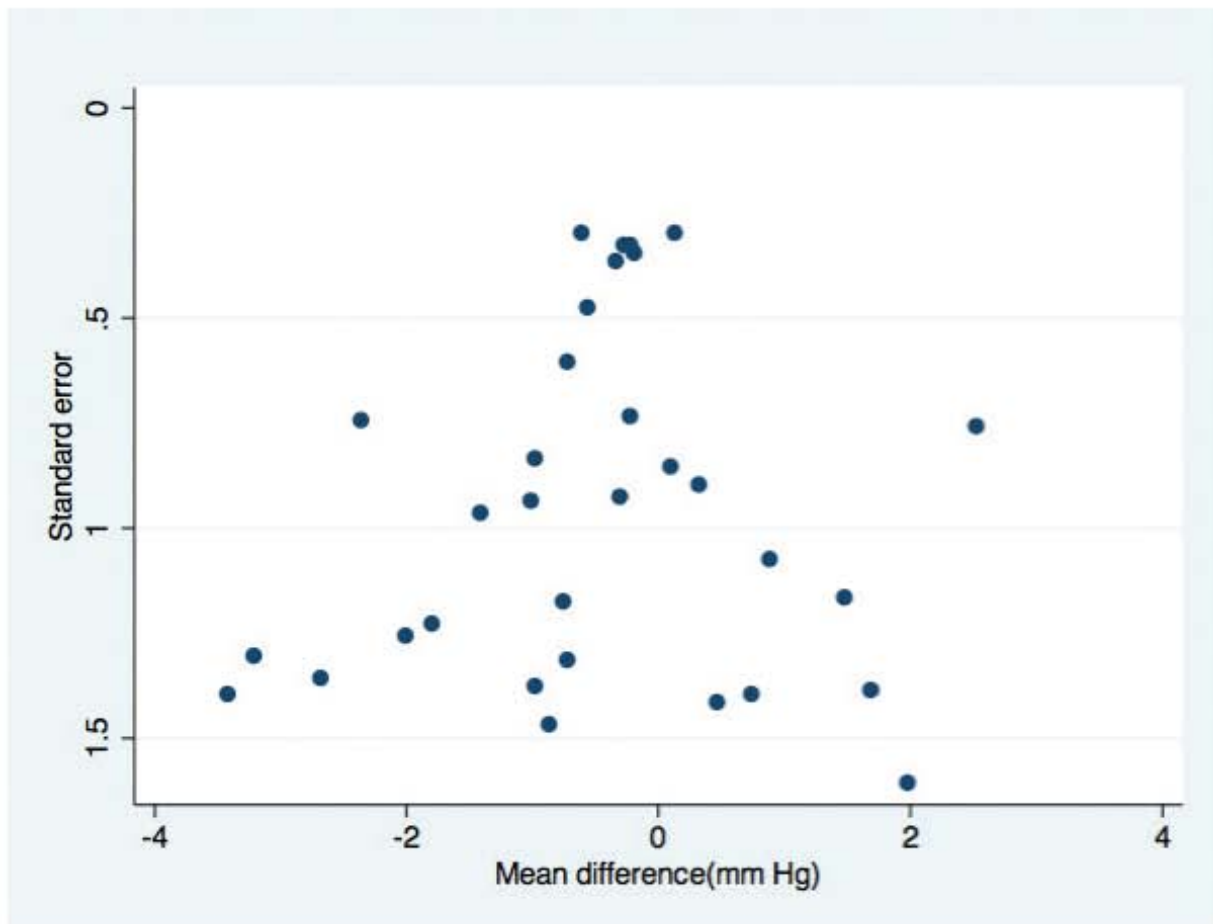


FIGURE 6.4

Funnel plot showing mean difference in diastolic blood pressure (mm Hg) by standard error of mean difference



References

- 1 Boulanger JM, Hill MD. Hypertension and stroke: 2005 Canadian Hypertension Educational Program recommendations. *Can J Neurol Sci* 2005; 32: 403–8.
- 2 Gluckman PD, Hanson MA, Cooper C, et al. Effect of in utero and early-life conditions on adult health and disease. *N Engl J Med* 2008; 359: 61–73.
- 3 Andersen LG, Angquist L, Eriksson JG, et al. Birth weight, childhood body mass index and risk of coronary heart disease in adults: combined historical cohort studies. *PLoS One*; 5: e14126.
- 4 Fomon S. Infant feeding in the 20th century: formula and beikost. *J Nutr* 2001; 131: 409S–20S.
- 5 Brunner EJ, Rees K, Ward K, et al. Dietary advice for reducing cardiovascular risk. *Cochrane Database Syst Rev* 2007: CD002128.
- 6 Singhal A, Cole TJ, Lucas A. Early nutrition in preterm infants and later blood pressure: two cohorts after randomised trials. *Lancet* 2001; 357: 413–9.
- 7 Geleijnse JM, Hofman A, Witteman JC, et al. Long-term effects of neonatal sodium restriction on blood pressure. *Hypertension* 1997; 29: 913–7.
- 8 Koletzko B, Agostoni C, Carlson SE, et al. Long chain polyunsaturated fatty acids (LC-PUFA) and perinatal development. *Acta Paediatr* 2001; 90: 460–4.
- 9 Morris MC, Sacks F, Rosner B. Does fish oil lower blood pressure? A meta-analysis of controlled trials. *Circulation* 1993; 88: 523–33.
- 10 Forsyth JS, Willatts P, Agostoni C, et al. Long chain polyunsaturated fatty acid supplementation in infant formula and blood pressure in later childhood: follow up of a randomised controlled trial. *BMJ* 2003; 326: 953.
- 11 de Jong C, Boehm G, Kikkert HK, et al. The Groningen LCPUFA study: No effect of short-term postnatal long-chain polyunsaturated fatty acids in healthy term infants on cardiovascular and anthropometric development at 9 years. *Pediatr Res*; 70: 411–6.
- 12 Owen CG, Martin RM, Whincup PH, et al. Effect of infant feeding on the risk of obesity across the life course: a quantitative review of published evidence. *Pediatrics* 2005; 115: 1367–77.
- 13 Colao A, Di Somma C, Cascella T, et al. Relationships between serum IGF1 levels, blood pressure, and glucose tolerance: an observational, exploratory study in 404 subjects. *Eur J Endocrinol* 2008; 159: 389–97.
- 14 Martin RM, Holly JM, Smith GD, et al. Could associations between breastfeeding and insulin-like growth factors underlie associations of breastfeeding with adult chronic disease? The Avon Longitudinal Study of Parents and Children. *Clin Endocrinol (Oxf)* 2005; 62: 728–37.
- 15 Jarvisalo MJ, Hutri-Kahonen N, Juonala M, et al. Breast feeding in infancy and arterial endothelial function later in life. The Cardiovascular Risk in Young Finns Study. *Eur J Clin Nutr* 2009; 63: 640–5.
- 16 Holmes VA, Cardwell C, McKinley MC, et al. Association between breast-feeding and anthropometry and CVD risk factor status in adolescence and young adulthood: the Young Hearts Project, Northern Ireland. *Public Health Nutr*; 13: 771–8.
- 17 de Jonge LL, van Osch-Gevers L, Geelhoed JJ, et al. Breastfeeding is not associated with left cardiac structures and blood pressure during the first two years of life. The Generation R Study. *Early Hum Dev*; 86: 463–8.
- 18 Naghettini AV, Belem JM, Salgado CM, et al. [Evaluation of risk and protection factors associated with high blood pressure in children]. *Arq Bras Cardiol*; 94: 486–91.
- 19 Fall CH, Borja JB, Osmond C, et al. Infant-feeding patterns and cardiovascular risk factors in young adulthood: data from five cohorts in low- and middle-income countries. *Int J Epidemiol*; 40: 47–62.
- 20 Rudnicka AR, Owen CG, Strachan DP. The effect of breastfeeding on cardiorespiratory risk factors in adult life. *Pediatrics* 2007; 119: e1107–15.
- 21 Brion MJ, Lawlor DA, Matijasevich A, et al. What are the causal effects of breastfeeding on IQ, obesity and blood pressure? Evidence from comparing high-income with middle-income cohorts. *Int J Epidemiol*; 40: 670–80.
- 22 Martin RM, Ness AR, Gunnell D, et al. Does breast-feeding in infancy lower blood pressure in childhood? The Avon Longitudinal Study of Parents and Children (ALSPAC). *Circulation* 2004; 109: 1259–66.
- 23 Evelein AM, Geerts CC, Visseren FL, et al. The association between breastfeeding and the cardiovascular system in early childhood. *Am J Clin Nutr*; 93: 712–8.
- 24 Khan F, Green FC, Forsyth JS, et al. The beneficial effects of breastfeeding on microvascular function in 11- to 14-year-old children. *Vasc Med* 2009; 14: 137–42.

- 25 Wilson AC, Forsyth JS, Greene SA, et al. Relation of infant diet to childhood health: seven year follow up of cohort of children in Dundee infant feeding study. *BMJ* 1998; 316: 21–5.
- 26 Kramer MS, Matush L, Vanilovich I, et al. Effects of prolonged and exclusive breastfeeding on child height, weight, adiposity, and blood pressure at age 6.5 y: evidence from a large randomized trial. *Am J Clin Nutr* 2007; 86: 1717–21.
- 27 Boulton J. Nutrition in childhood and its relationships to early somatic growth, body fat, blood pressure, and physical fitness. *Acta Paediatr Scand Suppl* 1981; 284: 1–85.
- 28 Zeman J, Simkova M. [Blood pressure values in infants and young children in relation to the duration of breast feeding]. *Cesk Pediatr* 1981; 36: 593–4.
- 29 Baranowski T, Bryan GT, Harrison JA, et al. Height, infant-feeding practices and cardiovascular functioning among 3 or 4 year old children in three ethnic groups. *J Clin Epidemiol* 1992; 45: 513–8.
- 30 Lawlor DA, Najman JM, Sterne J, et al. Associations of parental, birth, and early life characteristics with systolic blood pressure at 5 years of age: findings from the Mater-University study of pregnancy and its outcomes. *Circulation* 2004; 110: 2417–23.
- 31 Whincup PH, Cook DG, Shaper AG. Early influences on blood pressure: a study of children aged 5–7 years. *BMJ* 1989; 299: 587–91.
- 32 Williams S, St George IM, Silva PA. Intrauterine growth retardation and blood pressure at age seven and eighteen. *J Clin Epidemiol* 1992; 45: 1257–63.
- 33 Lucas A, Morley R. Does early nutrition in infants born before term programme later blood pressure? *BMJ* 1994; 309: 304–8.
- 34 Rona RJ, Qureshi S, Chinn S. Factors related to total cholesterol and blood pressure in British 9 year olds. *J Epidemiol Community Health* 1996; 50: 512–18.
- 35 Esposito-Del Puente A, Scalfi L, De Filippo E, et al. Familial and environmental influences on body composition and body fat distribution in childhood in southern Italy. *Int J Obes Relat Metab Disord* 1994; 18: 596–601.
- 36 Butler NR, Golding J, Haslum M, et al. Recent findings from the 1970 child health and education study: preliminary communication. *J R Soc Med* 1982; 75: 781–4.
- 37 Lawlor DA, Riddoch CJ, Page AS, et al. Infant feeding and components of the metabolic syndrome: findings from the European Youth Heart Study. *Arch Dis Child* 2005; 90: 582–8.
- 38 Owen CG, Whincup PH, Gilg JA, et al. Effect of breast feeding in infancy on blood pressure in later life: systematic review and meta-analysis. *BMJ* 2003; 327: 1189–95.
- 39 Taittonen L, Nuutinen M, Turtinen J, et al. Prenatal and postnatal factors in predicting later blood pressure among children: cardiovascular risk in young Finns. *Pediatr Res* 1996; 40: 627–32.
- 40 Kolacek S, Kapetanovic T, Luzar V. Early determinants of cardiovascular risk factors in adults. B. Blood pressure. *Acta Paediatr* 1993; 82: 377–82.
- 41 Leeson CP, Kattenhorn M, Deanfield JE, et al. Duration of breast feeding and arterial distensibility in early adult life: population based study. *BMJ* 2001; 322: 643–7.
- 42 Martin RM, McCarthy A, Smith GD, et al. Infant nutrition and blood pressure in early adulthood: the Barry Caerphilly Growth study. *Am J Clin Nutr* 2003; 77: 1489–97.
- 43 Ravelli AC, van der Meulen JH, Osmond C, et al. Infant feeding and adult glucose tolerance, lipid profile, blood pressure, and obesity. *Arch Dis Child* 2000; 82: 248–52.
- 44 Martin RM, Ben-Shlomo Y, Gunnell D, et al. Breast feeding and cardiovascular disease risk factors, incidence, and mortality: the Caerphilly study. *J Epidemiol Community Health* 2005; 59: 121–9.
- 45 Martin RM, Ebrahim S, Griffin M, et al. Breastfeeding and atherosclerosis: intima-media thickness and plaques at 65-year follow-up of the Boyd Orr cohort. *Arterioscler Thromb Vasc Biol* 2005; 25: 1482–8.
- 46 Wadsworth ME. Follow-up of the first national birth cohort: findings from the Medical Research Council National Survey of Health and Development. *Paediatr Perinat Epidemiol* 1987; 1: 95–117.
- 47 Fall CH, Osmond C, Barker DJ, et al. Fetal and infant growth and cardiovascular risk factors in women. *BMJ* 1995; 310: 428–32.

Total cholesterol

Ischemic heart disease is the leading cause of death worldwide (1), and blood lipids levels constitute important risk factors (2). It has been suggested that the development of heart disease may be programmed by exposures in early life (3), including breastfeeding practices.

Biological plausibility

The high cholesterol content of breastmilk may have a long-term programming effect on blood cholesterol levels. Higher intakes of cholesterol in infancy down-regulate hepatic hydroxymethylglutaryl coenzyme A (HMG-CoA), reducing the synthesis of cholesterol (4). Animal studies have reported that early exposure to high levels of cholesterol is negatively associated with later cholesterol levels. Furthermore, Devlin et al (5) observed that HMG-CoA reductase was higher in formula-fed than in milk-fed piglets, whereas LDL receptor mRNA was independent of early diet. HMG-CoA is the rate-limiting enzyme in synthesis of cholesterol from acetate, and HMG-CoA reductase inhibitors (statins) have a cholesterol lowering effect (6).

Overview of the evidence

Our 2007 systematic review and meta-analysis assessed the long-term consequences of breastfeeding duration, with total cholesterol being one of the outcomes. In that review, subjects who had been breastfed presented lower total cholesterol in adulthood [mean difference: -0.18 mmol/L; 95% confidence interval: -0.30; -0.06 mmol/L], whereas among children and adolescents no such association was observed. Publication bias was not present for studies assessing the long-term effect of breastfeeding on total cholesterol.

During the review process, we identified another meta-analysis that examined the effect of breastfeeding on blood cholesterol in adult life. Owen et al (7) systematically reviewed the evidence on the association between initial feeding (incidence of breastfeeding) and total cholesterol in adulthood (age > 16 years). Mean total cholesterol levels were slightly lower among subjects who were ever breastfed [mean difference: -0.04 mmol/L (95% confidence interval: -0.08; 0.00)].

Update of the 2007 meta-analysis

In the present update, the search strategy was the same as in the systematic review published in 2007. Below, we describe the additional eight studies that were identified in our new search.

Khan et al (8) evaluated subjects aged 11 to 14 years who had been followed in the first years of life as part of the Dundee Infant Feeding Study. Mean total cholesterol was similar among ever (mean 4.05 mmol/L) and never (mean 4.01 mmol/L) breastfed subjects.

Victora et al (30) evaluated subjects who have been followed up since birth in 1982, in Pelotas, a southern Brazilian city. At 18 years of age, mean total cholesterol was slightly higher among those male subjects who had been breastfed [mean difference 0.07 mmol/L (standard error: 0.06)].

Holmes et al (9) reported that breastfeeding was not related to blood cholesterol among subjects aged between 20 and 25 years who have been followed since adolescence. Information on breastfeeding was collected retrospectively at ages 12 and 15 years.

Williams et al (33) evaluated subjects who have been followed up since birth in 1972–73, in Dunedin, New Zealand. At 26 years of age, mean total cholesterol of women who had ever been breastfed was lower [mean difference -0.14 (standard error 0.08)] than that observed among the never breastfed women.

Järvisalo et al (10) assessed endothelial function and cardiovascular risk factors among subjects aged between 24 and 39 years who had been followed for 21 years. The manuscript provided crude estimates only, and total cholesterol was not associated with breastfeeding, in either sex.

In the Framingham Third Generation Study, Parikh et al (11) reported that breastfeeding was not associated with total cholesterol at mean age of 41 years.

Rudnicka et al (12) evaluated subjects who had been followed since birth in 1958, in England, Wales and Scotland (the British Birth Cohort). At 44–45 years of age, breastfeeding for > 1 month was not associated with total cholesterol.

Gunnarsdottir et al (13) also reported that breastfeeding was not associated with total cholesterol in adulthood, in a retrospective cohort in Iceland, information on infant feeding of 3614 subjects, who were born from 1914–1935, was gathered from midwives' birth records.

Our updated meta-analysis included 35 studies that provided 42 estimates on the effect of breastfeeding on total cholesterol (**Table 7.1**). **Figure 7.1** shows that there was marked heterogeneity among studies, with respect to the long-term effect of breastfeeding on total cholesterol. Indeed, the chi-square test for heterogeneity was statistically significant ($p=0.004$), and a random-effect model was used to pool the estimates. The mean difference was -0.01 mmol/L (95% confidence interval: -0.05; 0.02), suggesting no association between breastfeeding and total cholesterol levels.

In the 2007 meta-analysis, we observed that the effect of breastfeeding on mean cholesterol was modified by age at assessment, with significant protection restricted to adults. **Table 7.2** shows that in this updated meta-analysis, the effect modification by age was not so marked, and there was no significant interaction, nor was the benefit of breastfeeding among adults [mean difference: -0.03 (95% confidence interval: -0.08; 0.01)].

Other subgroup analyses were carried out. Length of recall ≥ 3 years resulted in larger, but non-significant differences [mean difference: -0.06 (95% confidence interval: -0.14; 0.02)]. On the other hand, control for confounding and for current body mass index did not modify the association between cholesterol and breastfeeding. In the meta-regression, age at assessment of blood cholesterol was the only variable whose inclusion in the model decreased heterogeneity among the studies, by 90.8%.

The funnel plot shows that small studies tended to report results in both directions, either negative or positive, suggesting no evidence of publication bias. Indeed, Egger's test was not statistically significant ($p=0.49$). Furthermore **Table 7.2** shows that the effect of breastfeeding on cholesterol was not related to study size.

When the analyses were restricted to the better designed studies (sample size of at least 700 individuals, including control of confounding factors), the pooled effect was equal to 0.00 mmol/L (95% CI -0.02;0.02).

Conclusion

In this updated meta-analysis, the beneficial effect of breastfeeding on total blood cholesterol in adulthood was smaller than that estimated by the earlier review. Compared to the original finding of -0.18 mmol/L in favor of breastfed subjects, the inclusion of new studies resulted in an estimate of -0.03 mmol/L (95% confidence interval: -0.08; 0.01). Because the confidence interval included the null effect, these results do not support a long-term programming effect of breastfeeding on blood lipids.

TABLE 7.1

Breastfeeding and total cholesterol in later life: studies included in the meta-analysis in ascending order of subjects' age at which outcome was measured

Author, Year	Year of birth of subjects	Study design	Mean age at total cholesterol assessment	Gender	Comparison groups	Mean difference in total cholesterol (standard error) in mmol / L among breastfed subjects
Jooste, 1991 (14)	1981–1986	Cohort	1 year	All	Exclusively breastfed vs. Formula fed for the first 3 months	-0.07 (SE 0.09)
Mize, 1995 (15)	Not stated	Cohort	1 year	All	Ever breastfed vs. Never breastfed	0.35 (SE 0.21)
Friedman, 1975 (16)	Not stated	Cross-sectional	2 years	All	Breastfed vs. Bottle fed	0.02 (SE 0.19)
Ward, 1980 (17)	1974	Cross-sectional	2 years	All	Breastfed for ≥ 1 month vs. Bottle fed	0.25 (SE 0.13)
Freedman, 1992 (18)	Not stated	Cross-sectional	3 years	All	Ever breastfed vs. Bottle fed	0.03 (SE 0.02)
Routi, 1997 (19)	1990–1992	Cohort	3 years	All	Breastmilk only source of milk at 7 months vs. Formula only source of milk at 7 months	0.12 (SE 0.06)
Huttunen, 1983 (20)	1975	Cohort	5 years	All	Breastfed for ≥ 6 months and not receiving formula vs. Formula fed between 1 and 6 months	0.2 (SE 0.11)
Crawford, 1981 (21)	1969–1970	Cohort	6 years	All	Exclusively BF vs. Exclusively FF	-0.04 (SE 0.20)
Elaraby, 1985 (22)	Not stated	Cross-sectional	7 years	All	Breastfed for > 6 months vs. Bottle fed	-0.57 (SE 0.21)
Plancoulaine, 2000 (23)	1981–1987	Cross-sectional	8 years	M F	Breastfed vs. Formula fed Breastfed vs. Formula fed	-0.3 (SE 0.13) 0 (SE 0.16)
Fomon, 1984 (24)	1966–1971	Cohort	8 years	M F	Breastfed vs. Formula fed Breastfed vs. Formula fed	-0.13 (SE 0.09) 0.03 (SE 0.12)
Hodgson, 1976 (25)	1962–1965	Cohort	9 years	M F	Exclusively breastfed vs. Formula fed for the first 3 months Exclusively breastfed vs. Formula fed for the first 3 months	0.54 (SE 0.34) 0.23 (SE 0.27)
Khan, 2009 (8)	1983–1986	Cohort	12 years	A	Breastfed vs. Not breastfed	0.04 (SE 0.12)
Hromodova, 1997 (26)	Not stated	Cohort	13 years	A	Breastfed vs. Formula fed	-0.09 (SE 0.27)
Owen, 2002 (27)	1982–1986	Cross-sectional	14 years	A	Exclusively breastfed vs Exclusively formula for the first 3 months	-0.02 (SE 0.04)
Singhal, 2004 (28)	1982–1985	RCT	14 years	A	Allocated to bank breastmilk vs. Allocated to preterm formula	-0.3 (SE 0.13)
Bergstrom, 1995 (29)	Not stated	Cross-sectional	15 years	A	Ever breastfed vs. Never breastfed	-0.03 (SE 0.10)

Author, Year	Year of birth of subjects	Study design	Mean age at total cholesterol assessment	Gender	Comparison groups	Mean difference in total cholesterol (standard error) in mmol / L among breastfed subjects
Friedman, 1975 (16)	Not stated	Cross-sectional	17 years	A	Breastfed vs. Bottle fed	0.03 (SE 0.25)
Victora, 2006 (30)	1982	Cohort	18 years	M	Ever breastfed vs. Never breastfed	0.07 (SE 0.06)
Kolacek, 1993 (31)	1968–1969	Cohort	20 years	M	Exclusively breastfed vs. Formula fed	-0.2 (SE 0.32)
				F	Exclusively breastfed vs. Formula fed	-0.2 (SE 0.31)
Holmes, 2010 (9)	1974–1978	Cohort	22 years	A	Breastfed vs. Not breastfed	0.2 (SE 0.13)
Leeson, 2001 (32)	1969–1975	Cross-sectional	24 years	A	Ever breastfed vs. Never breastfed	-0.18 (SE 0.11)
Williams, 2006 (33)	1972–1973	Cohort	26 years	F	Ever breastfed vs. Never breastfed	-0.14 (SE 0.08)
				M	Breastfed vs. Formula fed	-0.2 (SE 0.15)
				F	Breastfed vs. Formula fed	-0.2 (SE 0.11)
				M	Exclusive breastfed for 5 month vs. Never breastfed	-0.1 (SE 0.25)
Järvisalo, 2009 (10)	1962–1977	Cohort	31 years	F	Exclusive breastfed for 5 months vs. Never breastfed	-0.5 (SE 0.23)
				M	Exclusive breastfed for 5 months vs. Never breastfed	-0.5 (SE 0.23)
Marmot, 1980 (34)	1946	Cohort	31 years	F	Exclusive breastfed for 5 months vs. Never breastfed	-0.5 (SE 0.23)
Parikh, 2009 (11)	Not stated	Cohort	41 years	A	Breastfed vs. Not breastfed	-0.03 (SE 0.03)
Rudnicka, 2007 (12)	1958	Cohort	44 years	A	Breastfed > 1 month vs. Never breastfed	-0.01 (SE 0.02)
Ravelli, 2000 (35)	1943–1947	Cohort	50 years	A	Exclusively breastfed vs. Bottle-fed during stay in hospital	-0.16 (SE 0.11)
Lamont, 1998 (36)	1947	Cohort	50 years	A	Ever breastfed vs. Never breastfed	0.54 (SE 0.39)
Martin, 2005 (37)	1920–1934	Cohort	52 years	M	Breastfed vs. Bottle fed	0.04 (SE 0.07)
Gunnarsdottir, 2007 (13)	1914–1935	Cohort	55 years	M	Breastfed vs. Bottle fed	0 (SE 0.1)
				F	Breastfed vs. Bottle fed	0.1 (SE 0.13)
Huxley, 2004 (38)	1942–1944	Cohort	55 years	A	Ever breastfed vs. Never breastfed	0.33 (SE 0.46)
Fall, 1992 (39)	1920–1930	Cohort	65 years	M	Exclusively breastfed vs. Formula fed	-0.32 (SE 0.25)
Fall, 1995 (40)	1920–1930	Cohort	66 years	F	Exclusively breastfed vs. Formula fed	0.12 (SE 0.43)
Martin, 2005 (41)	1918–1939	Cohort	71 years	A	Breastfed vs. Bottle fed	0.12 (SE 0.14)

TABLE 7.2

Breastfeeding and blood cholesterol in later life. Random-effects meta-analyses of cholesterol levels by subgroup

Subgroup analysis	Number of estimates of total cholesterol	Mean difference (95% confidence interval)	P-value
By age group			
1 to 9 years	15	0.02 (-0.06; 0.11)	0.63
10 to 19 years	7	-0.01 (-0.08; 0.06)	0.73
≥ 20 years	20	-0.03 (-0.08; 0.01)	0.16
By study size			
< 200 participants	16	0.02 (-0.12; 0.16)	0.76
200–699 participants	15	-0.06 (-0.14; 0.03)	0.18
≥ 700 participants	11	0.00 (-0.02; -0.02)	0.98
By year at birth			
Before 1980	26	-0.01 (-0.07; 0.04)	0.64
After 1980	8	-0.03 (-0.12; 0.06)	0.57
Birth cohort			
Yes	20	-0.01 (-0.07; 0.06)	0.84
No	22	-0.02 (-0.07; 0.02)	0.50
By length of recall of breastfeeding			
< 3 years	31	0.01 (-0.03; 0.05)	0.71
≥ 3 years	11	-0.06 (-0.14; 0.02)	0.12
By categorization of breastfeeding			
Ever breastfed	26	-0.01 (-0.05; 0.02)	0.48
Breastfed for a given number of months	2	-0.02 (-0.13; 0.08)	0.68
By control for confounders			
Yes	14	-0.01 (-0.05; 0.03)	0.61
No	28	-0.02 (-0.09; 0.06)	0.61
By control for current measure of body size			
Yes	9	-0.03 (-0.11; 0.05)	0.45
No	33	-0.01 (-0.06; 0.04)	0.71
Total	42	-0.01 (-0.05; 0.02)	

Note: The total number of studies does not always add to 42 due missing information on year of birth (8 studies), categorization of breastfeeding (14 studies)

FIGURE 7.1

Mean difference in total cholesterol in mmol/L (and its 95% confidence interval) between breastfed and non-breastfed subjects in different studies. Whether the estimate was for males (M), females (F) and all (A) is indicated in parenthesis.

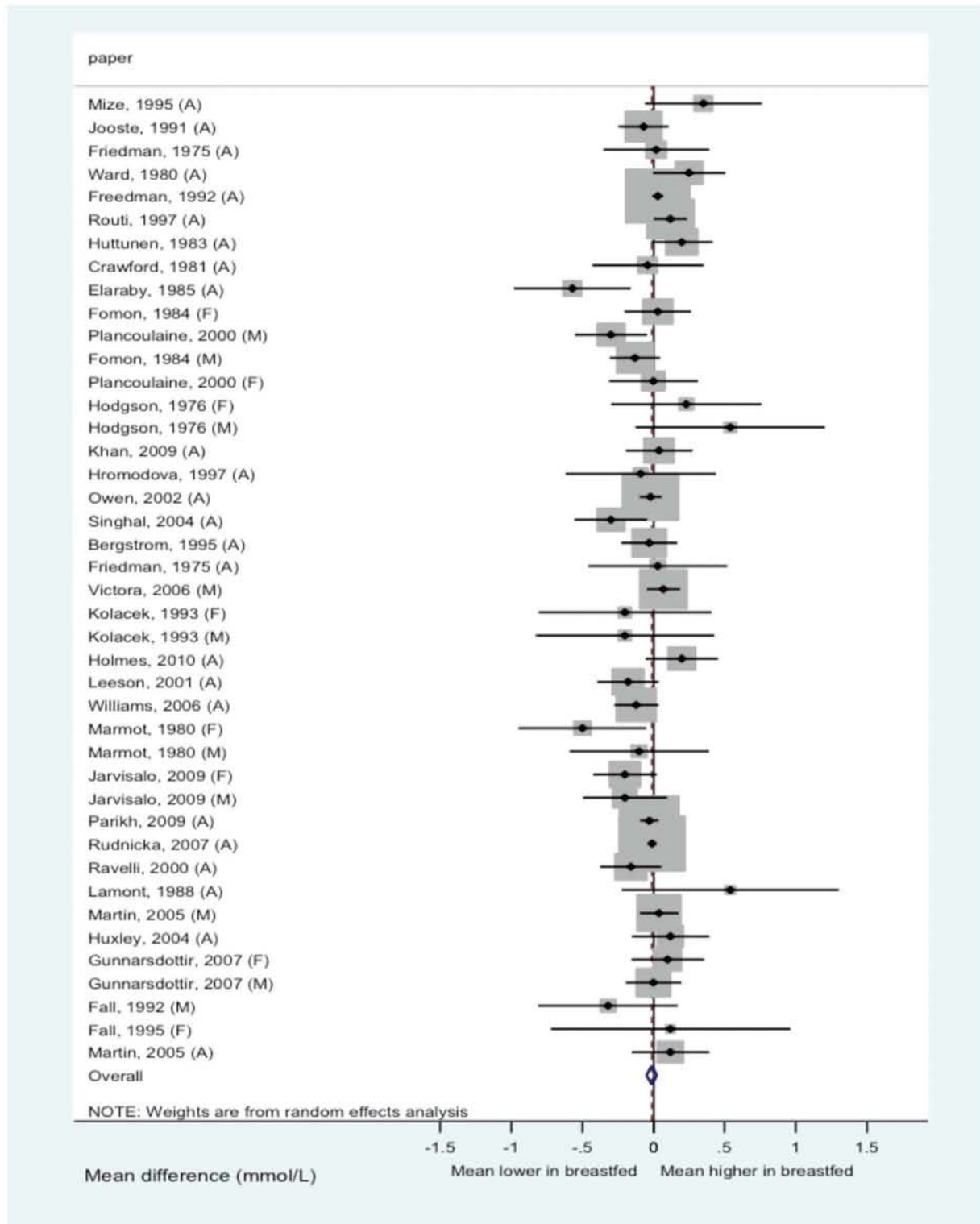
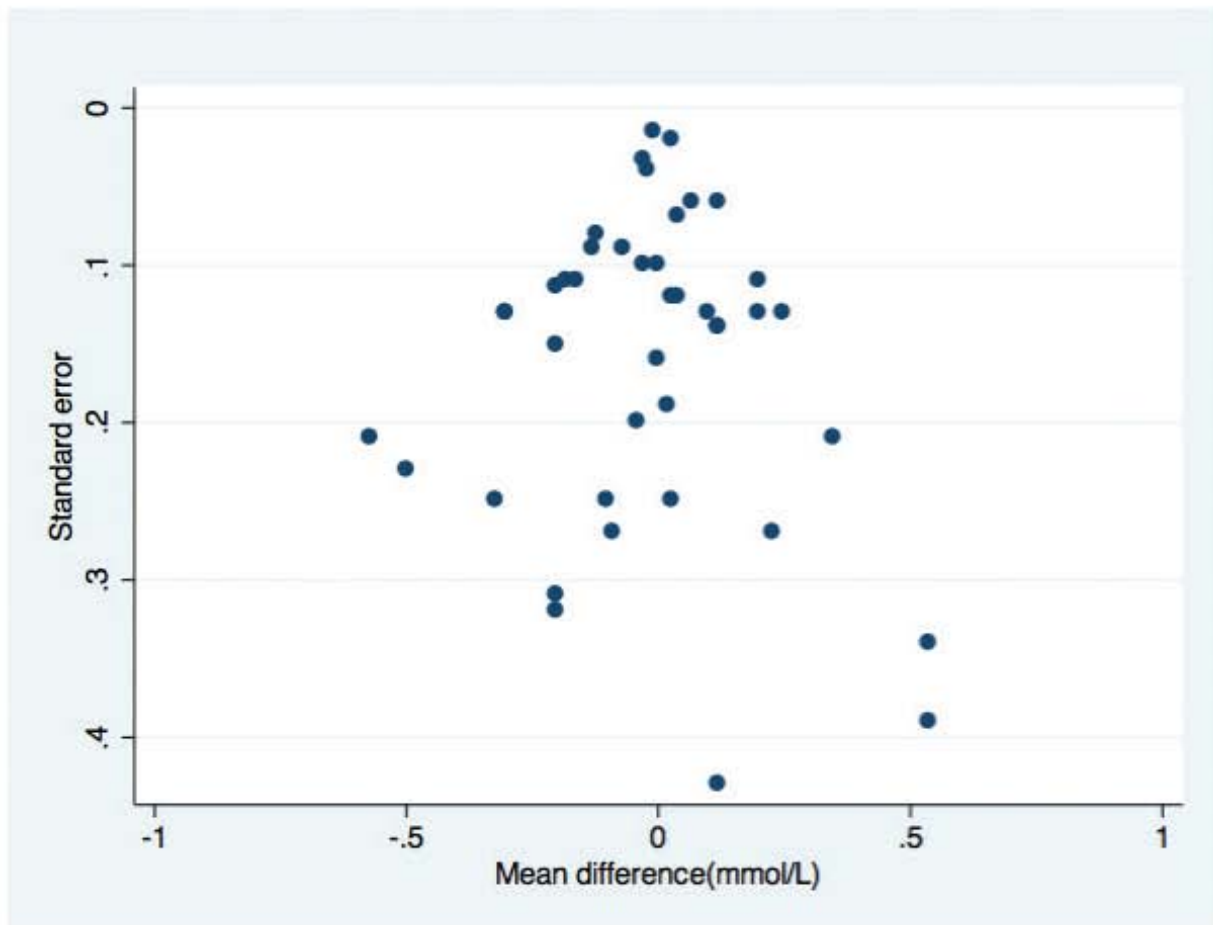


FIGURE 7.2

Funnel plot showing mean difference in total cholesterol (mmol/L) by standard error of mean difference



References

- 1 Lopez AD, Mathers CD, Ezzati M, et al. Global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet* 2006; 367: 1747–57.
- 2 Wilson PW, D'Agostino RB, Levy D, et al. Prediction of coronary heart disease using risk factor categories. *Circulation* 1998; 97: 1837–47.
- 3 Lauren L, Jarvelin MR, Elliott P, et al. Relationship between birthweight and blood lipid concentrations in later life: evidence from the existing literature. *Int J Epidemiol* 2003; 32: 862–76.
- 4 Wong WW, Hachey DL, Insull W, et al. Effect of dietary cholesterol on cholesterol synthesis in breast-fed and formula-fed infants. *J Lipid Res* 1993; 34: 1403–11.
- 5 Devlin AM, Innis SM, Shukin R, et al. Early diet influences hepatic hydroxymethyl glutaryl coenzyme A reductase and 7 α -hydroxylase mRNA but not low-density lipoprotein receptor mRNA during development. *Metabolism* 1998; 47: 20–6.
- 6 Baigent C, Keech A, Kearney PM, et al. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90,056 participants in 14 randomised trials of statins. *Lancet* 2005; 366: 1267–78.
- 7 Owen CG, Whincup PH, Kaye SJ, et al. Does initial breastfeeding lead to lower blood cholesterol in adult life? A quantitative review of the evidence. *Am J Clin Nutr* 2008; 88: 305–14.
- 8 Khan F, Green FC, Forsyth JS, et al. The beneficial effects of breastfeeding on microvascular function in 11- to 14-year-old children. *Vasc Med* 2009; 14: 137–42.
- 9 Holmes VA, Cardwell C, McKinley MC, et al. Association between breast-feeding and anthropometry and CVD risk factor status in adolescence and young adulthood: the Young Hearts Project, Northern Ireland. *Public Health Nutr*; 13: 771–8.
- 10 Jarvisalo MJ, Hutri-Kahonen N, Juonala M, et al. Breast feeding in infancy and arterial endothelial function later in life. The Cardiovascular Risk in Young Finns Study. *Eur J Clin Nutr* 2009; 63: 640–5.
- 11 Parikh NI, Hwang SJ, Ingelsson E, et al. Breastfeeding in infancy and adult cardiovascular disease risk factors. *Am J Med* 2009; 122: 656–63 e1.
- 12 Rudnicka AR, Owen CG, Strachan DP. The effect of breastfeeding on cardiorespiratory risk factors in adult life. *Pediatrics* 2007; 119: e1107–15.
- 13 Gunnarsdottir I, Aspelund T, Birgisdottir BE, et al. Infant feeding patterns and midlife erythrocyte sedimentation rate. *Acta Paediatr* 2007; 96: 852–6.
- 14 Jooste PL, Rossouw LJ, Steenkamp HJ, et al. Effect of breast feeding on the plasma cholesterol and growth of infants. *J Pediatr Gastroenterol Nutr* 1991; 13: 139–42.
- 15 Mize CE, Uauy R, Kramer R, et al. Lipoprotein-cholesterol responses in healthy infants fed defined diets from ages 1 to 12 months: comparison of diets predominant in oleic acid versus linoleic acid, with parallel observations in infants fed a human milk-based diet. *J Lipid Res* 1995; 36: 1178–87.
- 16 Friedman G, Goldberg SJ. Concurrent and subsequent serum cholesterol of breast- and formula-fed infants. *Am J Clin Nutr* 1975; 28: 42–5.
- 17 Ward SD, Melin JR, Lloyd FP, et al. Determinants of plasma cholesterol in children – a family study. *Am J Clin Nutr* 1980; 33: 63–70.
- 18 Freedman DS, Lee SL, Byers T, et al. Serum cholesterol levels in a multiracial sample of 7,439 preschool children from Arizona. *Prev Med* 1992; 21: 162–76.
- 19 Routi T, Ronnema T, Viikari JS, et al. Tracking of serum lipoprotein (a) concentration and its contribution to serum cholesterol values in children from 7 to 36 months of age in the STRIP Baby Study. Special Turku Coronary Risk Factor Intervention Project for Babies. *Ann Med* 1997; 29: 541–7.
- 20 Huttunen JK, Saarinen UM, Kostinen E, et al. Fat composition of the infant diet does not influence subsequent serum lipid levels in man. *Atherosclerosis* 1983; 46: 87–94.
- 21 Crawford PB, Clark MJ, Pearson DL, et al. Serum cholesterol of 6-year-olds in relation to environmental factors. *J Am Diet Assoc* 1981; 78: 41–6.
- 22 Elaraby I, Aref G. Effect of breast vs. humanized milk on concurrent and subsequent serum cholesterol levels. *Saudi Medical Journal* 1985; 6: 480–3.
- 23 Plancoulaine S, Charles MA, Lafay L, et al. Infant-feeding patterns are related to blood cholesterol concentration in prepubertal children aged 5–11 y: the Fleurbaix-Laventie Ville Sante study. *Eur J Clin Nutr* 2000; 54: 114–9.

- 24 Fomon SJ, Rogers RR, Ziegler EE, et al. Indices of fatness and serum cholesterol at age eight years in relation to feeding and growth during early infancy. *Pediatr Res* 1984; 18: 1233–8.
- 25 Hodgson PA, Ellefson RD, Elveback LR, et al. Comparison of serum cholesterol in children fed high, moderate, or low cholesterol milk diets during neonatal period. *Metabolism* 1976; 25: 739–46.
- 26 Hromadova M, Kostalova L, Leskova L, et al. Relationship between the duration of the breast-feeding period and the lipoprotein profile of children at the age of 13 years. *Physiol Res* 1997; 46: 21–5.
- 27 Owen CG, Whincup PH, Odoki K, et al. Infant feeding and blood cholesterol: a study in adolescents and a systematic review. *Pediatrics* 2002; 110: 597–608.
- 28 Singhal A, Cole TJ, Fewtrell M, et al. Breastmilk feeding and lipoprotein profile in adolescents born preterm: follow-up of a prospective randomised study. *Lancet* 2004; 363: 1571–8.
- 29 Bergstrom E, Hernell O, Persson LA, et al. Serum lipid values in adolescents are related to family history, infant feeding, and physical growth. *Atherosclerosis* 1995; 117: 1–13.
- 30 Victora CG, Horta BL, Post P, et al. Breast feeding and blood lipid concentrations in male Brazilian adolescents. *J Epidemiol Community Health* 2006; 60: 621–5.
- 31 Kolacek S, Kapetanovic T, Zimolo A, et al. Early determinants of cardiovascular risk factors in adults. A. Plasma lipids. *Acta Paediatr* 1993; 82: 699–704.
- 32 Leeson CP, Kattenhorn M, Deanfield JE, et al. Duration of breast feeding and arterial distensibility in early adult life: population based study. *BMJ* 2001; 322: 643–7.
- 33 Williams MJ, Williams SM, Poulton R. Breast feeding is related to C reactive protein concentration in adult women. *J Epidemiol Community Health* 2006; 60: 146–8.
- 34 Marmot MG, Page CM, Atkins E, et al. Effect of breast-feeding on plasma cholesterol and weight in young adults. *J Epidemiol Community Health* 1980; 34: 164–7.
- 35 Ravelli AC, van der Meulen JH, Osmond C, et al. Infant feeding and adult glucose tolerance, lipid profile, blood pressure, and obesity. *Arch Dis Child* 2000; 82: 248–52.
- 36 Lamont DW, Parker L, Cohen MA, et al. Early life and later determinants of adult disease: a 50 year follow-up study of the Newcastle Thousand Families cohort. *Public Health* 1998; 112: 85–93.
- 37 Martin RM, Ben-Shlomo Y, Gunnell D, et al. Breast feeding and cardiovascular disease risk factors, incidence, and mortality: the Caerphilly study. *J Epidemiol Community Health* 2005; 59: 121–9.
- 38 Huxley RR, Neil HA. Does maternal nutrition in pregnancy and birth weight influence levels of CHD risk factors in adult life? *Br J Nutr* 2004; 91: 459–68.
- 39 Fall CH, Barker DJ, Osmond C, et al. Relation of infant feeding to adult serum cholesterol concentration and death from ischaemic heart disease. *BMJ* 1992; 304: 801–5.
- 40 Fall CH, Osmond C, Barker DJ, et al. Fetal and infant growth and cardiovascular risk factors in women. *BMJ* 1995; 310: 428–32.
- 41 Martin RM, Ebrahim S, Griffin M, et al. Breastfeeding and atherosclerosis: intima-media thickness and plaques at 65-year follow-up of the Boyd Orr cohort. *Arterioscler Thromb Vasc Biol* 2005; 25: 1482–8.

Type-2 diabetes

Type-2 diabetes is the most common form of diabetes, being related to an increased risk of cardiovascular diseases (1). The incidence of type-2 diabetes has increased worldwide and it is also becoming more frequent among children and adolescents (2). The epidemic of type-2 diabetes has been partly attributed to an increase in obesity and a decrease in physical activity. There is also evidence that type-2 diabetes may be programmed by diet in early life (3).

Biological plausibility

At least three biological pathways have been proposed for explaining a protective effect of breastfeeding on type-2 diabetes.

The presence of long-chain polyunsaturated fatty acids (LCPUFAs) in breastmilk, including docosahexanoic acid (DHA) and arachidonic acid (AA), has been proposed as a possible mechanism for the effect of breastfeeding on the incidence of type-2 diabetes. Breastfeeding increases LCPUFAs in skeletal muscle membrane, which in turn is inversely related to fasting glucose. Therefore, early changes in skeletal muscle membrane due to LCPUFAs saturation may protect against insulin resistance, β -cell failure, and type-2 diabetes (4).

Differences in insulin secretion according to infant feeding constitute another potential mechanism for the long-term programming of breastfeeding on glucose metabolism. Formula-fed infants have higher concentrations of insulin, which may lead to β -cell failure and type-2 diabetes (5).

It has also been suggested that breastfeeding protects against obesity, and this may constitute another pathway to type-2 diabetes (6). However, it is unlikely that this is an important pathway, because – as discussed in [Chapter 5](#) – current evidence suggests that breastfeeding has a small protective effect against overweight or obesity.

Overview of the evidence

In 2007, we carried out a systematic review and meta-analysis that assessed the long-term consequences of breastfeeding duration, in which type-2 diabetes was one of the outcomes. In that meta-analysis, breastfeeding presented a protective effect against type-2 diabetes (pooled odds ratio: 0.63; 95% CI: 0.45; 0.89).

During that review process, we identified a 2006 meta-analysis that examined the effect of breastfeeding on type-2 diabetes, and concentration of insulin and glucose. Owen et al (3) reported that breastfeeding was associated with a lower risk of type-2 diabetes (pooled odds ratio: 0.61; 95% CI: 0.44; 0.85).

Update of the 2007 meta-analysis

In the present update, the search strategy was the same as in the systematic review published in 2007. Three additional studies were identified, which are described below.

Mayer-Davis et al (7) carried a case-control study on type-2 diabetes including 10–21 year-old cases from six clinical sites across South Carolina and Colorado, using controls from primary care clinics in the same area. After adjustment for several confounders, the odds ratio for type-2 diabetes among those subjects who were ever breastfed was 0.43 (95% CI: 0.19; 0.99). Adjustment for current BMI z-score reduced the strength of the association (0.82 [95% CI; 0.30; 2.30]).

Parikh et al (8) evaluated the protective effect of breastfeeding for cardiovascular disease risk factors among subjects enrolled in the Framingham Third Generation cohort. The mean age of the studied population was 41 years, and diabetes was defined as fasting plasma glucose ≥ 126 mg/dL or treatment with either insulin or oral hypoglycemic agents. The odds ratio of diabetes among ever breastfed subjects was 0.40 (95% CI: 0.09; 1.70), after adjustment for several confounding variables.

Fall et al (9) analyzed data from the COHORTS collaboration (Consortium on Health Orientated Research in Transitional Societies) comprising birth-cohort studies in five low and middle income countries: Brazil, Guatemala, India, Philippines and South Africa. The adjusted analyses showed that diabetes (defined as a glucose concentration ≥ 7.0 mmol/l) was not related to ever breastfeeding [odds ratio: 1.25 (95% CI: 0.63; 2.51)]. It should be noted, however, that very few infants were never breastfed in these cohorts. Additional analyses showed a U-shaped association between breastfeeding duration and diabetes, with the lowest risk in the intermediate duration categories, but the association was not statistically significant.

Our updated meta-analysis included 10 studies (**Table 8.1**). **Figure 8.1** shows that there was heterogeneity among studies ($p=0.02$). In a random-effects model, subjects who were breastfed were less likely to present type 2-diabetes, with the pooled odds ratio equal to 0.66 (95% CI: 0.49; 0.89).

Table 8.2 shows that the protective effect of breastfeeding was higher among studies involving adolescents, compared to adults. The effect of breastfeeding was smaller in studies that adjusted for body mass index (pooled odds ratio: 0.79), suggesting that obesity may partly mediate the association. The small number of studies precluded the statistical assessment of the main sources of heterogeneity.

Interpretation of funnel plot is challenging, given the small number of studies. Nevertheless, **Figure 8.2** suggests that there is publication bias. In spite of one small study showing a higher risk of diabetes among breastfed subjects, small studies reporting moderate odds of diabetes for breastfed subjects seem to be missing.

Due to the small number of studies, it was not possible to select a sizeable subgroup with stronger designs. Two large studies (8,9) had samples greater than 500 subjects and controlled for confounding factors, but their odds ratios were very different, 1.25 and 0.40 respectively. The latter adjusted for current body weight as well as socioeconomic position, which may have underestimated the true protection. In light of such heterogeneity, no firm conclusions can be drawn from the two studies with the strongest designs.

Conclusion

The evidence suggests that breastfeeding may have a protective effect against type-2 diabetes, particularly among adolescents. Obesity/overweight seems to account for part of the association. Generalization from these findings is restricted by the small number of studies and the presence of significant heterogeneity among them.

TABLE 8.1

Breastfeeding and type-2 diabetes in later life: studies included in ascending order of subjects' age at which outcome was measured

Author, Year	Study design	Mean age at assessment	Gender	Categorization of breastfeeding	Odds ratio of type-2 diabetes among breastfed subjects
Young, 2002 (10)	Case-control	13 years	All	Breastfed $\geq$ 6 months vs. Breastfed < 6 months	0.36 (0.13; 0.99)
Mayer-Davis, 2008 (7)	Case-control	15 years	All	Ever breastfed vs. Never breastfed	0.82 (0.30; 2.30)
Evenhouse, 2005 (11)	Cross sectional	15 years	All	Ever breastfed vs. Never breastfed	0.40 (SE: 0.24)
Fall, 2011 (9)	Cohort	25 years	All	Ever breastfed vs. Never breastfed	1.25 (0.63; 2.51)
Petit, 1997 (12)	Cohort	25 years	All	Exclusively breastfed vs. exclusively bottle fed	0.41 (0.18; 0.93)
Rich-Edwards, 2004 (13)	Cohort	59 years	Female	Ever breastfed vs. Never breastfed	0.79 (0.74; 0.85)
Parikh, 2009 (8)	Cohort	41 years	All	Ever breastfed vs. Never breastfed	0.40 (0.09; 1.70)
Ravelli, 2000 (14)	Cohort	50 years	All	Exclusively breastfed vs. bottle-fed during stay in hospital	0.51 (0.3; 0.9)
Martin, 2005 (15)	Cross-sectional	52 years	Male	Ever breastfed vs. Never breastfed	2.89 (0.65; 12.83)
Martin, 2005 (16)	Cohort	71 years	All	Ever breastfed vs. Never breastfed	0.97 (0.41; 2.30)

TABLE 8.2

Breastfeeding and the risk of type-2 diabetes in later life: Random-effects meta-analyses of risk of type-2 diabetes by subgroup.

Subgroup analysis	Number of estimates	Pooled odds ratio and 95% confidence interval	P Value
By age group			
10 to 19 years	3	0.44 (0.30; 0.65)	< 0.001
≥ 20 years	7	0.76 (0.55; 1.04)	0.09
By study size			
< 500 participants	3	0.54 (0.29; 0.99)	0.04
≥ 500 participants	6	0.70 (0.47; 1.04)	0.08
By study design			
Cohort	6	0.73 (0.55; 0.96)	0.02
Cross-sectional or case-control	4	0.63 (0.30; 1.31)	0.22
By adjustment for BMI			
No	7	0.64 (0.44; 0.93)	0.02
Yes	3	0.79 (0.43; 1.44)	0.44
Total	10	0.66 (0.49; 0.89)	

FIGURE 8.1

Odds ratio and 95% confidence interval of having type-2 diabetes in different studies, comparing breastfed vs. non-breastfed subjects.

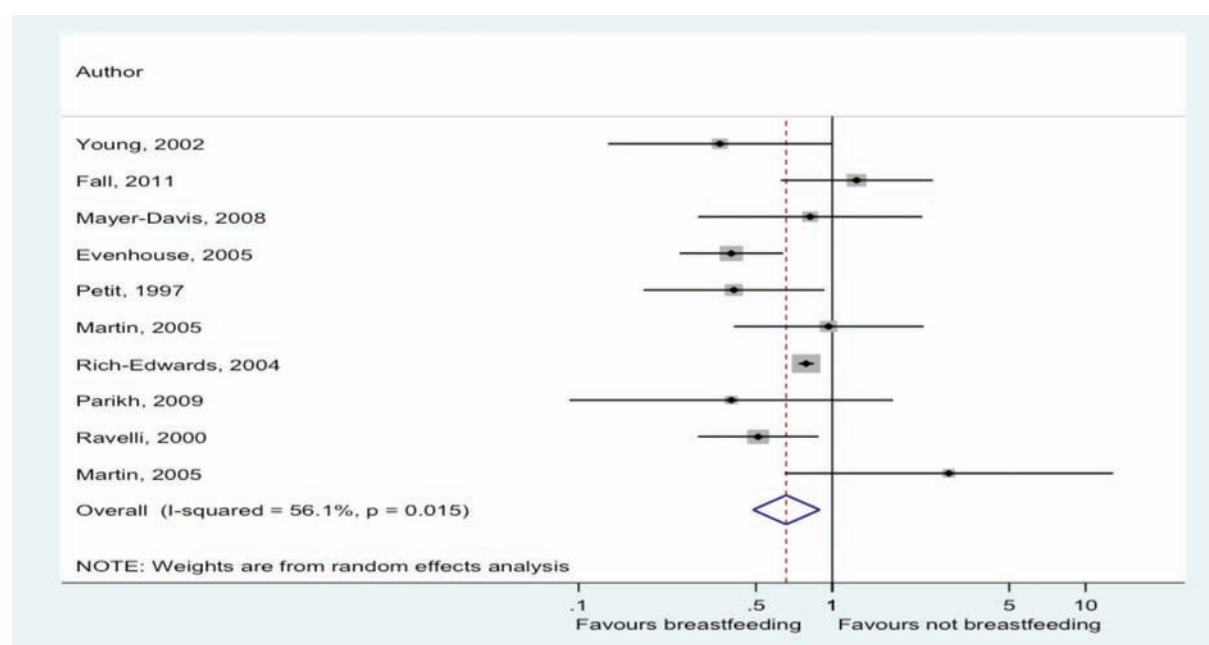
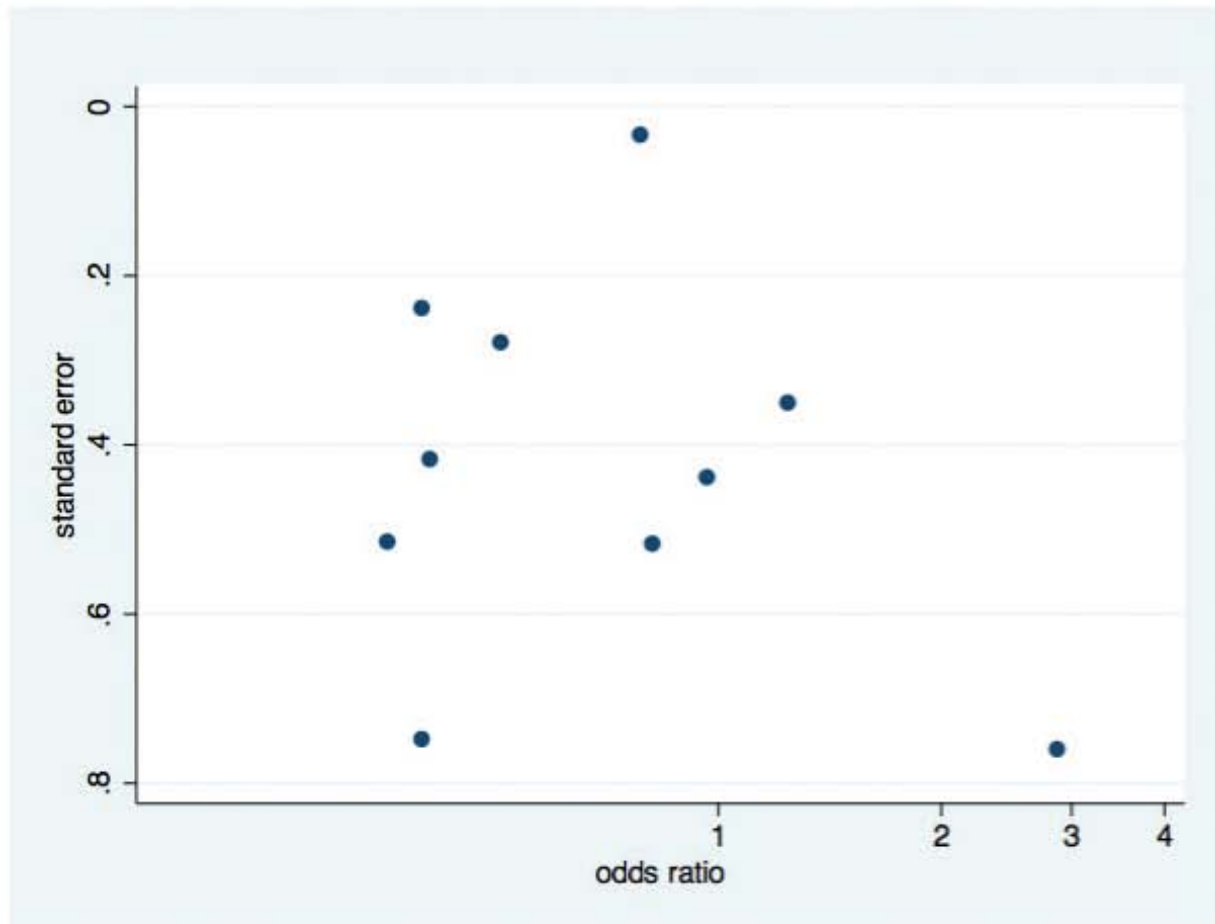


FIGURE 8.2

Funnel plot showing odds ratio for type-2 diabetes by standard error of odds ratio



References

- 1 Vinicor F. Is diabetes a public-health disorder? *Diabetes Care* 1994; 17 Suppl 1: 22–7.
- 2 Cohen MS. Fetal and childhood onset of adult cardiovascular diseases. *Pediatr Clin North Am* 2004; 51: 1697–719, x.
- 3 Owen CG, Martin RM, Whincup PH, et al. Does breastfeeding influence risk of type 2 diabetes in later life? A quantitative analysis of published evidence. *Am J Clin Nutr* 2006; 84: 1043–54.
- 4 Baur LA, O'Connor J, Pan DA, et al. The fatty acid composition of skeletal muscle membrane phospholipid: its relationship with the type of feeding and plasma glucose levels in young children. *Metabolism* 1998; 47: 106–12.
- 5 Salmenpera L, Perheentupa J, Siimes MA, et al. Effects of feeding regimen on blood glucose levels and plasma concentrations of pancreatic hormones and gut regulatory peptides at 9 months of age: comparison between infants fed with milk formula and infants exclusively breast-fed from birth. *J Pediatr Gastroenterol Nutr* 1988; 7: 651–6.
- 6 Owen CG, Martin RM, Whincup PH, et al. Effect of infant feeding on the risk of obesity across the life course: a quantitative review of published evidence. *Pediatrics* 2005; 115: 1367–77.
- 7 Mayer-Davis EJ, Dabelea D, Lamichhane AP, et al. Breast-feeding and type 2 diabetes in the youth of three ethnic groups: the SEARCH for diabetes in youth case-control study. *Diabetes Care* 2008; 31: 470–5.
- 8 Parikh NI, Hwang SJ, Ingelsson E, et al. Breastfeeding in infancy and adult cardiovascular disease risk factors. *Am J Med* 2009; 122: 656–63 e1.
- 9 Fall CH, Borja JB, Osmond C, et al. Infant-feeding patterns and cardiovascular risk factors in young adulthood: data from five cohorts in low- and middle-income countries. *Int J Epidemiol*; 40: 47–62.
- 10 Young TK, Martens PJ, Taback SP, et al. Type 2 diabetes mellitus in children: prenatal and early infancy risk factors among native Canadians. *Arch Pediatr Adolesc Med* 2002; 156: 651–5.
- 11 Evenhouse E, Reilly S. Improved estimates of the benefits of breastfeeding using sibling comparisons to reduce selection bias. *Health Serv Res* 2005; 40: 1781–802.
- 12 Pettitt DJ, Forman MR, Hanson RL, et al. Breastfeeding and incidence of non-insulin-dependent diabetes mellitus in Pima Indians. *Lancet* 1997; 350: 166–8.
- 13 Rich-Edwards JW, Stampfer MJ, Manson JE, et al. Breastfeeding during infancy and the risk of cardiovascular disease in adulthood. *Epidemiology* 2004; 15: 550–6.
- 14 Ravelli AC, van der Meulen JH, Osmond C, et al. Infant feeding and adult glucose tolerance, lipid profile, blood pressure, and obesity. *Arch Dis Child* 2000; 82: 248–52.
- 15 Martin RM, Ben-Shlomo Y, Gunnell D, et al. Breast feeding and cardiovascular disease risk factors, incidence, and mortality: the Caerphilly study. *J Epidemiol Community Health* 2005; 59: 121–9.
- 16 Martin RM, Ebrahim S, Griffin M, et al. Breastfeeding and atherosclerosis: intima-media thickness and plaques at 65-year follow-up of the Boyd Orr cohort. *Arterioscler Thromb Vasc Biol* 2005; 25: 1482–8.

Performance in intelligence tests

The relationship between breastfeeding and child development has been studied since long ago. In 1929, Hoefer and Hardy (1) observed that breastfeeding was positively associated with intelligence among children aged from 7 to 13 years. In 1950, Douglas reported that duration of breastfeeding was inversely related to the age at which the child started walking (2). These early findings were followed by several studies reporting that breastfeeding was positively associated with child development (3).

Biological Plausibility

Some studies have proposed that the benefits of breastfeeding in terms of child development may be due to improved care rather than better nutrition (4). Nevertheless, there are several biological mechanisms that may account for a nutritional effect.

The presence of long-chain polyunsaturated fatty acids in breastmilk, including docosahexanoic acid (DHA) and arachidonic acid (AA), is a possible mechanism for its effect on development (5). AA and DHA, which are important for retinal and cortical brain development (6,7), accumulate in the brain and retina during the last trimester of pregnancy and the first months after birth (8). Breastfed infants have higher concentrations of these fatty acids (9). Indeed, Isaacs et al reported that breastfeeding was positively related to brain volume and – among male subjects – white matter was also related to breastfeeding (10), thus suggesting that breast milk promotes structural changes in the brain.

Genetic epidemiology techniques are starting to be employed to explore the role played by DHA metabolism in the association between breastfeeding and intelligence (11). By using genetic variants as proxies for DHA exposure, confounding is avoided, because the inheritance of one trait is independent of socio-behavioral variables or of other genetic traits. Particular attention has been given to the FADS2 gene which affects the metabolism of DHA (12,13,14,15,16). However, evidence on how the FADS2 genotype may modify the effect of breastfeeding on intelligence tests is still controversial (17,18).

Breastfeeding is also positively associated with maternal sensitivity and with bonding between mother and infant (19), which constitute an additional mechanism for the long-term effect of breastfeeding on cognition.

Specific methodological issues

General methodological issues have already been addressed in the Introduction. When assessing the association between infant feeding and later performance in intelligence tests, two methodological issues deserve special attention: self-selection bias and residual confounding.

At least in high-income societies, breastfeeding mothers are likely to be more health-conscious, and, therefore, to promote healthy habits and provide stimulation to their infants. For example, Der et al (20) observed that breastfeeding mothers were more likely to stimulate their infants. Because stimulation at home is positively related to performance in intelligence tests (21,22), studies assessing the long-term consequences of infant feeding on cognition should adjust their estimates for variables measuring home stimulation. Reinforcing the importance of such adjustment, it has been reported that the magnitude of the association between breastfeeding and cognition is reduced after adjustment for home stimulation (20).

Residual confounding by socioeconomic status should also be considered when interpreting the evidence on breastfeeding and intelligence. Because performance in intellectual tests is higher among wealthy and educated subjects (23), residual confounding may affect the assessment of causality in settings where breastfeeding is also directly associated with socioeconomic status. Even large studies accounting for several potential confounders may still be affected, if such confounders were not properly measured or were mis-specified in the regression model. Comparison of observational studies with different confounding structures may be used to assess this possibility.

Overview of the evidence

In 2007, we carried out a systematic review and meta-analysis that assessed the long-term consequences of breastfeeding duration. Performance in intelligence tests was one of the outcomes we assessed. We employed the following selection criteria:

- cognition had to be assessed using standard tests;
- estimates had to be adjusted for stimulation or interaction with the child.

In that meta-analysis, breastfeeding was associated with higher scores in intelligence tests (mean difference: 4.9 points; 95% CI: 2.97; 6.92).

Update of the 2007 meta-analysis

In this updated review, eight additional studies were identified, which are described below.

Der et al (20) studied the children of women enrolled in the US National Longitudinal Survey of Youth (NLSY79), which has been following a cohort of young people aged 14 to 22 since 1979. Children with birthweight < 2500g or gestational age < 35 weeks were excluded. Data on 5475 children was obtained and IQ was assessed through the Peabody individual achievement test. Adjustment for stimulation at home, maternal IQ, socioeconomic and demographic variables, and birthweight reduced the difference in performance in intelligence tests from 4.69 points (standard error: 0.38) to 0.52 (standard error: 0.36).

Gibson-Davis et al (24) studied 1645 children enrolled in the Fragile Families and Child Wellbeing Study, a cohort of mostly unmarried parents and their children from 20 cities in 15 North American states. Maternal and child cognition were assessed during the visit at 3 years using the Peabody Picture Vocabulary Test-III (PPVT-III). In the crude analysis the PPVT-III score was 6.6 points higher among those children who were breastfed. After controlling for maternal performance in PPVT-III, stimulation at home, socioeconomic and demographic variables the difference between breastfed and non-breastfed children fell to 1.72 points (standard error: 0.60).

Clark et al (25) followed a cohort of healthy full-term children born in urban communities near Santiago, Chile and had been enrolled in a trial on prevention of iron deficiency. At five years, 784 children were evaluated and cognition was assessed using the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R). Those who were breastfed for more than 8 months scored 1.0 point (95% confidence interval: -1.1; 3.1) higher than those breastfed < 2 months, after adjustment for home stimulation, maternal IQ, socioeconomic variables, and nutritional status at 12 months. For children who were breastfed for 2 to 8 months, the score was 1.3 points higher than that for children breastfed < 2 months.

Evenhouse et al (26) used data from 2743 sibling pairs enrolled in the National Longitudinal Study of Adolescent Health, United States, to assess the relationship between breastfeeding and cognition. Of these, 288 pairs comprised one ever breastfed and one never breastfed sibling. IQ was evaluated using the PPVT test. Ever breastfed siblings scored on average 1.68 percentile point (standard error 0.94) higher. Sibling comparisons are one of strategies that can be used to improve causal inference, by controlling for confounders as well as for self-selection bias (assuming that maternal characteristics are shared among siblings). On the other hand, heterogeneity in breastfeeding duration is smaller than that for unrelated subjects; therefore, precision of estimates is reduced.

Zhou et al (4) studied 302 children from a trial aimed at assessing the efficacy of iron supplementation in pregnancy. IQ was assessed at 4 years of age using the Stanford-Binet Intelligence Scale. After adjustment for home stimulation and socioeconomic status, the IQ of children who were breastfed for at least 6 months was 0.8 (standard error 1.29) points higher than the observed among children who were breastfed for less than 6 months.

Eickmann et al (27) studied a cohort of 205 children born in a poor area in the interior of Pernambuco (Northeast Brazil). At 12 months, child development was assessed using the Bayley Scales of Infant Development II. After control for confounding (home stimulation, socioeconomic status and haemoglobin level), the mean score was 3.0 (95% confidence interval: 0.48; 5.53) points higher among those children who were exclusively or predominantly breastfed at 1 month.

Whitehouse et al (28) used data from the follow-up of the Western Australian Pregnancy (Raine) Cohort to assess whether breastfeeding was related to performance in the PPVT-R test at 10 years of age. In the multivariable regression, children who were predominantly breastfed for > 6 months scored 4.04 points higher than those who never breastfed, after adjustment for maternal age at conception, maternal smoking during pregnancy, alcohol intake during pregnancy, maternal schooling, parity, socioeconomic status, and home stimulation

In this new meta-analysis, we included 13 studies with 14 estimates on the effect of breastfeeding on performance in intelligence tests. When interpreting these results it is important to note that different tests for intellectual performance were included (although all of these had a reference mean value of 100 with a standard deviation of 15). In addition, each comparison included two groups according to the duration of breastfeeding, but the cut-off points were different from study to study (in 6 studies, the comparison was never vs ever breastfed; in the remainder different duration of breastfeeding were used as cut-off). **Figure 9.1** shows that there was marked heterogeneity among the studies. Because this was statistically significant, a random-effect model was used to pool the estimates. In the pooled analysis, breastfed subjects presented higher performance in intelligence tests (mean difference: 3.45 points; 95% confidence interval: 1.92; 4.98).

Maternal IQ is positively associated with offspring cognition as well as with breastfeeding duration, at least in high-income societies where most studies were carried out (29,30). We assessed whether adjustment or matching by maternal intelligence was a source of heterogeneity among the study findings. Studies were considered as having controlled for maternal IQ if (a) they included maternal IQ in the multivariable analysis, (b) were randomized clinical trials or (c) discordant sibling pair analyses. **Figure 9.2** shows that there was marked heterogeneity among studies, with smaller effects in studies that controlled for maternal IQ. Adjustment for maternal IQ explained 77.3% of the heterogeneity among the studies; the pooled effect from studies with control of maternal IQ was equal to 2.19 (95% CI: 0.89; 3.50).

Further adjustment for other study characteristics (length of recall for breastfeeding information, age at measurement of outcome, type of categorization of breastfeeding duration, control of confounding factors other than maternal IQ, type of design) did not provide further explanation for heterogeneity. **Figure 9.3** shows that year of publication was also related to heterogeneity among studies, and the reported benefit of breastfeeding was lower for those studies published from 2006 onwards. Adjustment for year of publication explained 51.9% of the heterogeneity.

The funnel plot was clearly asymmetrical, and small studies reporting a negative effect of breastfeeding on cognition appear to be missing (**Figure 9.4**). Nevertheless, the overall effect of breastfeeding on intelligence cannot be explained by such bias. **Table 9.1** shows that the mean effect was similar among those studies with sample sizes < 500 (mean effect: 3.61; 95% confidence interval: 1.59; 5.63) and ≥ 500 subjects (mean effect: 3.25; 95% confidence interval: 1.09; 5.40). **Table 9.1** also shows that categorization of breastfeeding duration did not modify the association between breastfeeding duration and performance in intelligence tests. On the other hand, those studies that were carried in subjects < 10 years of age observed a higher mean difference than studies involving older subjects.

Relevant studies not included in the present review

Two randomized trials were identified. The first, by Lucas et al (31), compared preterm infants who were allocated to breastmilk or to formula. Because this study entailed a comparison of infants who received different types of milk, it was included in the meta-analysis. The second was the cluster randomized trial on promotion of breastfeeding by Kramer et al (32) in Belarus, in which 31 maternity hospitals and surrounding clinics were included (16 in the intervention and 15 in the comparison group). The proportion of infants exclusively breastfed at 3 and 6 months was higher among infants from the intervention group and performance in the Wechsler Abbreviated Scale of Intelligence at 6.5 years of age was also higher in the intervention group, with cluster-mean IQ differences of 7.5 points (95% confidence interval: 0.8; 14.3). These analyses were based on intent to treat and are very relevant to this review because of the randomized design. Nevertheless, the quantitative estimates could not be included in the meta-analyses because both groups – intervention and comparison – included breastfed and non-breastfed subjects.

As mentioned in the methodology section, residual confounding by socioeconomic status should be taken into account when interpreting the evidence on association between breastfeeding and intelligence. Brion et al (33) compared the association between breastfeeding and cognition in two cohorts, one from a high-income country (ALSPAC, UK) and another from a middle-income country (Pelotas, Brazil). In ALSPAC socioeconomic status was positively associated with breastfeeding duration, whereas in Pelotas there was no such association. In both settings, duration of breastfeeding was directly related to cognition. The authors concluded that the observed association is unlikely to

be explained by residual socioeconomic confounding. These two datasets were not included in the meta-analyses because there was no adjustment for home stimulation.

Conclusion

This meta-analysis suggests that breastfeeding is associated with increased performance in intelligence tests in childhood and adolescence, of 3.5 points on average. Maternal IQ is an important confounder, but it accounts for only part of this association – even among those studies that adjusted for maternal intelligence, breastfeeding was associated with an additional 2.19 IQ points. The two existing randomized trial on this issue also reported significant benefits of breastfeeding, suggesting that this association is causal.

On the other hand, the practical implications of a small increase in the performance in intelligence tests may be open to debate. Nevertheless, one Brazilian study suggests that breastfeeding is associated with achieved schooling in adolescents, in a population where breastfeeding duration does not present marked variability by socioeconomic position (34).

TABLE 9.1

Breastfeeding and mean difference in cognitive development scores and its 95% confidence interval between breastfed and non-breastfed subjects in different studies mean: Random-effects meta-analysis by subgroups

Subgroup analysis	Number of estimates	Mean difference (95% confidence interval)	P Value
By age group			
1 to 9 years	8	4.74 (2.41; 7.08)	< 0.001
10 to 19 years	4	2.50 (0.97; 4.03)	0.001
By study size			
< 500 participants	7	3.61 (1.59; 5.63)	< 0.001
≥ 500 participants	7	3.25 (1.09; 5.40)	0.003
By categorization of breastfeeding			
Ever breastfed	5	3.25 (0.91; 5.58)	0.006
Breastfed for a given number of months	9	3.52 (1.53; 5.51)	0.001
Total	14	3.45 (1.92; 4.98)	

Note: The total number of studies does not add to 14, due to missing information on age at assessment of intellectual performance (2 studies).

FIGURE 9.1

Mean difference in cognitive development scores and its 95% confidence interval between breastfed and non-breastfed subjects in different studies. Whether the estimate was for males (M), females (F) and all (A) is indicated in parenthesis.

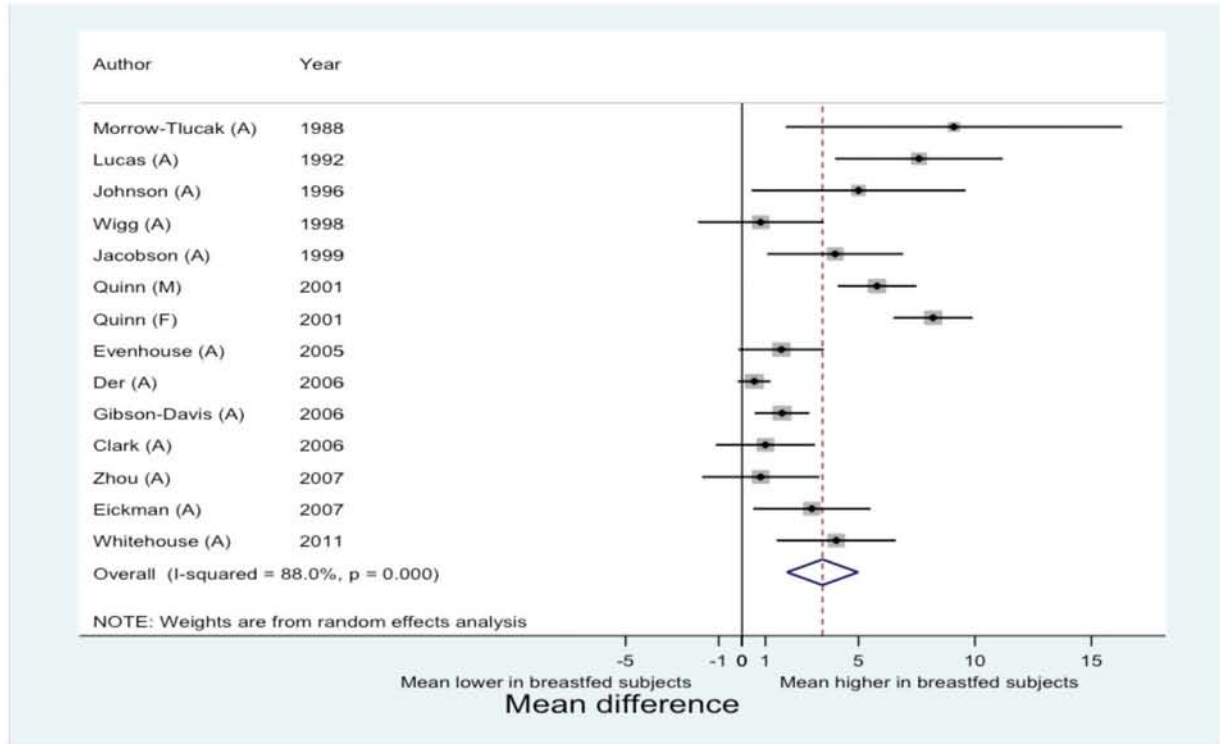


FIGURE 9.2

Mean difference in cognitive development scores and its 95% confidence interval between breastfed and non-breastfed subjects in different studies, stratified by adjustment for maternal IQ in analysis or in the study design

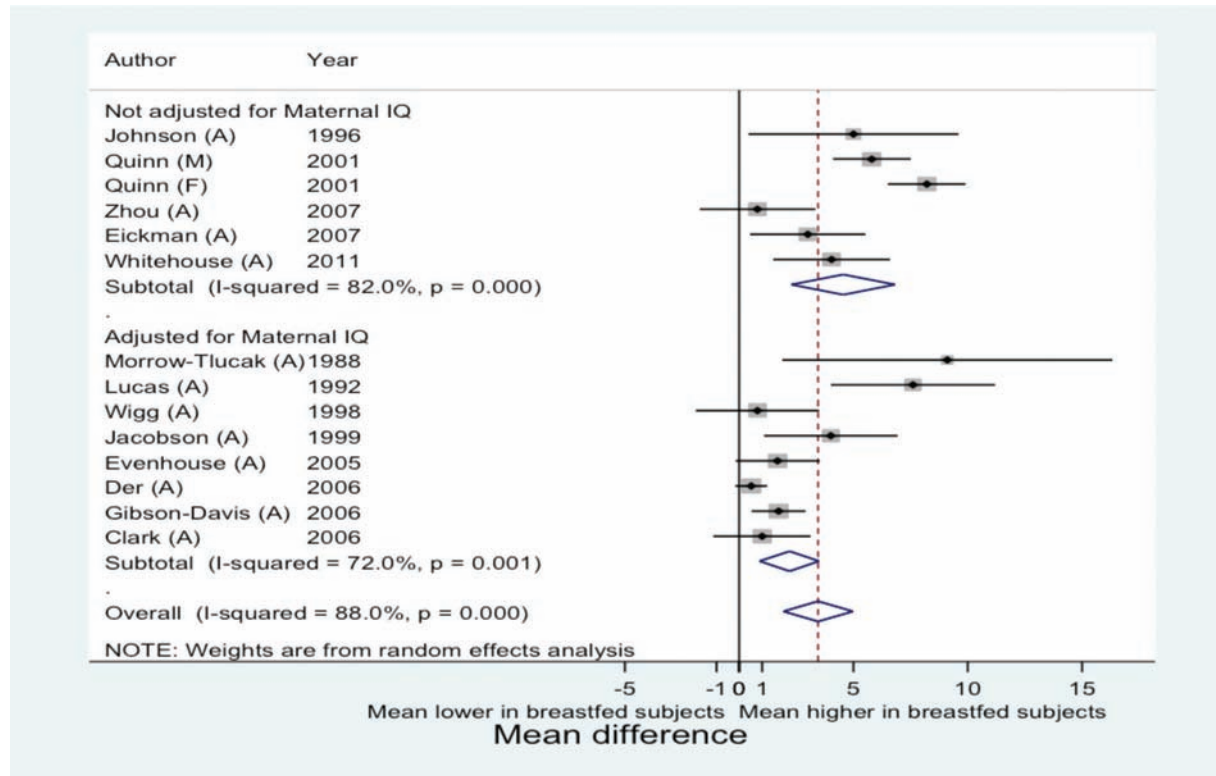


FIGURE 9.3

Mean difference in cognitive development scores and its 95% confidence interval between breastfed and non-breastfed subjects in different studies, stratified by year of publication

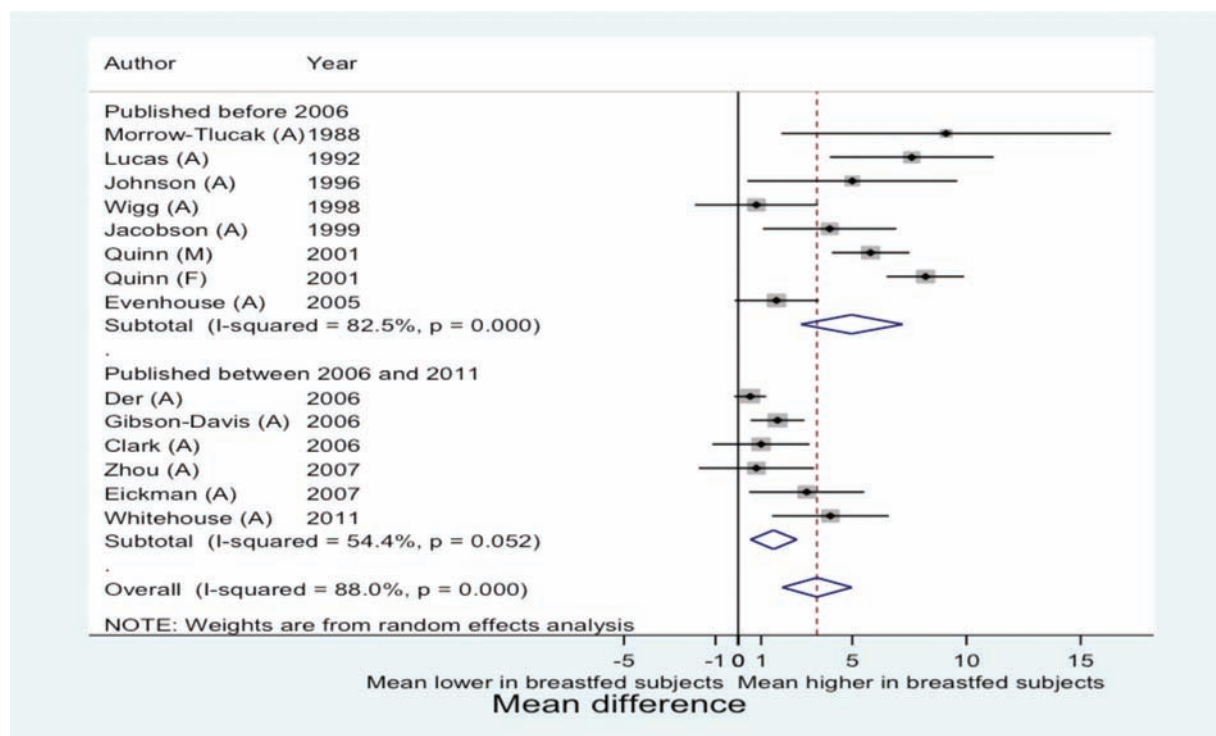
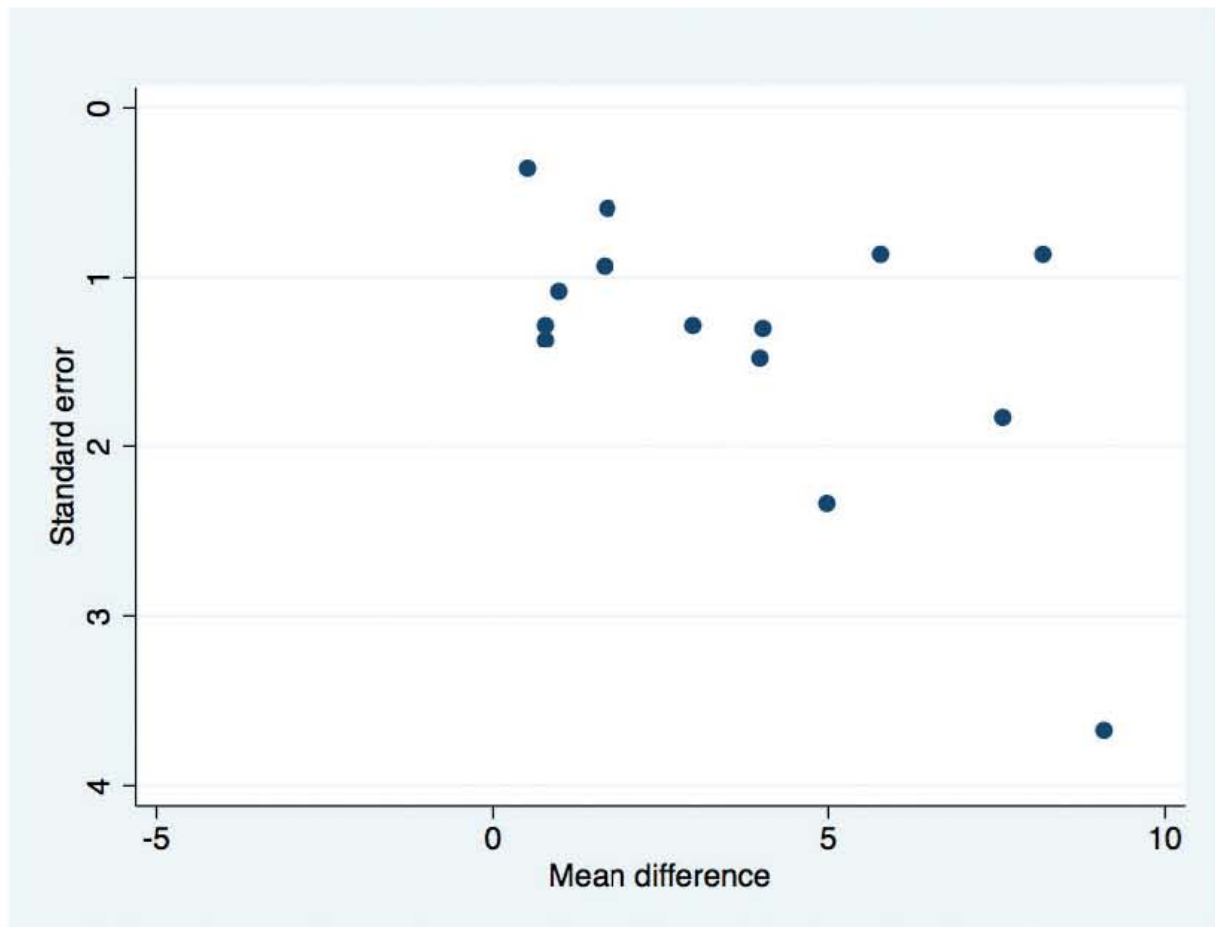


FIGURE 9.4

Funnel plot showing mean difference in performance in intelligence tests by standard error of mean difference



References

- 1 Hoefer C HMC. Later development of breast fed and artificially fed infants. Comparison of physical and mental growth. *JAMA* 1929; 92: 615–9.
- 2 Douglas J. Breast feeding in Great Britain in 1946. Part II. Infant feeding in relation to survival, health and development. *J Obstet Gynaecol Br Commonw* 1950; 57: 349–61.
- 3 Anderson JW, Johnstone BM, Remley DT. Breast-feeding and cognitive development: a meta-analysis. *Am J Clin Nutr* 1999; 70: 525–35.
- 4 Zhou SJ, Baghurst P, Gibson RA, et al. Home environment, not duration of breast-feeding, predicts intelligence quotient of children at four years. *Nutrition* 2007; 23: 236–41.
- 5 Koletzko B, Agostoni C, Carlson SE, et al. Long chain polyunsaturated fatty acids (LC-PUFA) and perinatal development. *Acta Paediatr* 2001; 90: 460–4.
- 6 Crawford MA. The role of essential fatty acids in neural development: implications for perinatal nutrition. *Am J Clin Nutr* 1993; 57: 703S–9S; discussion 9S–10S.
- 7 Birch EE, Birch DG, Hoffman DR, et al. Dietary essential fatty acid supply and visual acuity development. *Invest Ophthalmol Vis Sci* 1992; 33: 3242–53.
- 8 Clandinin MT, Chappell JE, Leong S, et al. Intrauterine fatty acid accretion rates in human brain: implications for fatty acid requirements. *Early Hum Dev* 1980; 4: 121–9.
- 9 Farquharson J, Cockburn F, Patrick WA, et al. Infant cerebral cortex phospholipid fatty-acid composition and diet. *Lancet* 1992; 340: 810–3.
- 10 Isaacs EB, Fischl BR, Quinn BT, et al. Impact of breast milk on intelligence quotient, brain size, and white matter development. *Pediatr Res* 67: 357–62.
- 11 Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol* 2003; 32: 1–22.
- 12 Cho HP, Nakamura MT, Clarke SD. Cloning, expression, and nutritional regulation of the mammalian Delta-6 desaturase. *J Biol Chem* 1999; 274: 471–7.
- 13 Park WJ, Kothapalli KS, Lawrence P, et al. An alternate pathway to long-chain polyunsaturates: the FADS2 gene product Delta8-desaturates 20:2n-6 and 20:3n-3. *J Lipid Res* 2009; 50: 1195–202.
- 14 Schaeffer L, Gohlke H, Muller M, et al. Common genetic variants of the FADS1 FADS2 gene cluster and their reconstructed haplotypes are associated with the fatty acid composition in phospholipids. *Hum Mol Genet* 2006; 15: 1745–56.
- 15 Tanaka T, Shen J, Abecasis GR, et al. Genome-wide association study of plasma polyunsaturated fatty acids in the InCHIANTI Study. *PLoS Genet* 2009; 5: e1000338.
- 16 Xie L, Innis SM. Genetic variants of the FADS1 FADS2 gene cluster are associated with altered (n-6) and (n-3) essential fatty acids in plasma and erythrocyte phospholipids in women during pregnancy and in breast milk during lactation. *J Nutr* 2008; 138: 2222–8.
- 17 Caspi A, Williams B, Kim-Cohen J, et al. Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism. *Proc Natl Acad Sci U S A* 2007; 104: 18860–5.
- 18 Steer CD, Davey Smith G, Emmett PM, et al. FADS2 polymorphisms modify the effect of breastfeeding on child IQ. *PLoS One*; 5: e11570.
- 19 Britton JR, Britton HL, Gronwaldt V. Breastfeeding, sensitivity, and attachment. *Pediatrics* 2006; 118: e1436–43.
- 20 Der G, Batty GD, Deary IJ. Effect of breast feeding on intelligence in children: prospective study, sibling pairs analysis, and meta-analysis. *BMJ* 2006; 333: 945.
- 21 Walker SP, Chang SM, Powell CA, et al. Effects of early childhood psychosocial stimulation and nutritional supplementation on cognition and education in growth-stunted Jamaican children: prospective cohort study. *Lancet* 2005; 366: 1804–7.
- 22 Barros AJ, Matijasevich A, Santos IS, et al. Child development in a birth cohort: effect of child stimulation is stronger in less educated mothers. *Int J Epidemiol*; 39: 285–94.
- 23 Bradley RH, Corwyn RF. Socioeconomic status and child development. *Annu Rev Psychol* 2002; 53: 371–99.
- 24 Gibson-Davis CM, Brooks-Gunn J. Breastfeeding and verbal ability of 3-year-olds in a multicity sample. *Pediatrics* 2006; 118: e1444–51.

- 25 Clark KM, Castillo M, Calatroni A, et al. Breast-feeding and mental and motor development at 51/2 years. *Ambul Pediatr* 2006; 6: 65–71.
- 26 Evenhouse E, Reilly S. Improved estimates of the benefits of breastfeeding using sibling comparisons to reduce selection bias. *Health Serv Res* 2005; 40: 1781–802.
- 27 Eickmann SH, de Lira PI, Lima Mde C, et al. Breast feeding and mental and motor development at 12 months in a low-income population in northeast Brazil. *Paediatr Perinat Epidemiol* 2007; 21: 129–37.
- 28 Whitehouse AJ, Robinson M, Li J, et al. Duration of breast feeding and language ability in middle childhood. *Paediatr Perinat Epidemiol*; 25: 44–52.
- 29 Bouchard TJ, Jr., McGue M. Genetic and environmental influences on human psychological differences. *J Neurobiol* 2003; 54: 4–45.
- 30 Silva PA, Buckfield P, Spears GF. Some maternal and child developmental characteristics associated with breast feeding: a report from the Dunedin Multidisciplinary Child Development Study. *Aust Paediatr J* 1978; 14: 265–8.
- 31 Lucas A, Morley R, Cole TJ, et al. Breast milk and subsequent intelligence quotient in children born preterm. *Lancet* 1992; 339: 261–4.
- 32 Kramer MS, Aboud F, Mironova E, et al. Breastfeeding and child cognitive development: new evidence from a large randomized trial. *Arch Gen Psychiatry* 2008; 65: 578–84.
- 33 Brion MJ, Lawlor DA, Matijasevich A, et al. What are the causal effects of breastfeeding on IQ, obesity and blood pressure? Evidence from comparing high-income with middle-income cohorts. *Int J Epidemiol*; 40: 670–80.
- 34 Victora CG, Barros FC, Horta BL, et al. Breastfeeding and school achievement in Brazilian adolescents. *Acta Paediatr* 2005; 94: 1656–60.

Conclusions

Relative to the 2007 systematic review, the present analyses rely on a considerably larger evidence base, including 60 additional studies published in the last five years. Availability of additional data has also allowed more detailed breakdowns of the studies, and a better understanding of reasons for heterogeneity. **Table 10.1** summarizes the effect measures from the meta-analyses.

TABLE 10.1

Pooled effects for each outcome, from all studies and from those deemed to be of higher scientific quality

Outcome	Pooled effect (95% confidence interval)	
	All studies	High-quality studies ^a
Mean total blood cholesterol (mmol/L)	-0.01 (-0.05; 0.02)	0.00 (-0.02; 0.02)
Mean systolic blood pressure (mmHg)	-1.02 (-1.45; -0.59)	-0.71 (-1.24; -0.19)
Mean diastolic blood pressure (mmHg)	-0.37 (-0.71; -0.04)	-0.27 (-0.64; 0.09)
Odds ratio of type-2 diabetes	0.66 (0.49-0.89)	Not estimated
Odds ratio of overweight/obesity	0.76 (0.71; 0.81)	0.88 (0.83; 0.93)
Mean performance in intelligence test (points)	3.45 (1.92-4.98)	2.19 (0.89-3.50)

^a High-quality studies include those with larger sample sizes and adjustment for confounding variables relevant to each outcome (see individual chapters for further details).

The meta-analyses of overweight/obesity, blood pressure, diabetes and intelligence suggest that benefits are larger for children and adolescents, and smallest among adults, suggesting a gradual dilution of the effect with time.

Because the meta-analyses are almost exclusively based on observational studies (see **Chapter 2**), the possibility of self-selection and residual confounding must be considered. Even when multiple studies show similar results, these studies may largely be subject to the same biases. In particular, most studies are derived from high-income settings where breastfeeding is more common among highly-educated, better-off mothers who are more health conscious, and whose offspring may be less likely to suffer from the negative outcomes covered in the present review. The fact that estimates from high-quality studies, with tighter control of confounding, are less marked than those from all studies suggests that this may be the case.

Interpretation of results from observational studies may be aided by also considering the findings of the only two randomized studies on the topic. The first is the follow up of English preterm infants allocated to breast or formula milk in early life, which showed a positive effect of breastfeeding on IQ, but no significant or sizeable effects on blood pressure or BMI (1,2). The second was the Belarus PRO-BIT trial whose mothers were allocated to breastfeeding promotion. These intent-to-treat analyses

compared an intervention group in which over 40% of subjects were exclusively breastfed at three months of age, with a control arm where only 6% were exclusively breastfed. Even though the true effects of breastfeeding may be underestimated, analyses at 7 years of age showed significant effects on IQ. No association was found with blood pressure or overweight/obesity. Results on cholesterol, diabetes or glucose levels are not available.

Interpreting the results of the present meta-analyses jointly with those of the two randomized trials, our conclusions are outlined below.

Total cholesterol. There was no effect in the overall meta-analyses. In the 2007 review, there was a significant effect among adults, which is no longer present in the updated analyses. The UK trial of preterm infants showed a small protective effect (3), while the Belarus trial did not report on this outcome. We conclude that breastfeeding does not seem to protect against total cholesterol levels.

Blood pressure. The pooled estimate from the high-quality studies indicates a small reduction of less than 1 mmHg in systolic pressure among breastfed subjects, and no significant protection in terms of diastolic pressure. Residual confounding may be an important problem. The Belarus and UK preterm trials found no effect of breastfeeding (2,4). We conclude that the protective effect of breastfeeding, if any, is too small to be of public health significance.

Diabetes. There was substantial protection in the pooled analyses, with a 34% reduction, but few studies are available and their results were considerably heterogeneous. Only two high-quality studies were identified, with conflicting results (one showing an increase and another a reduction among breastfed subjects). The randomized trials did not present any results on these outcomes. Our conclusion is that further studies are needed on this outcome.

Overweight-obesity. In the pooled analyses of all studies, breastfeeding was associated with a 24% reduction in overweight and/or obesity, but the reduction was only 12% in the high-quality studies. Residual confounding may be still affecting these results, because protection is not evident in studies from low and middle-income countries where the social patterning of breastfeeding is not clear cut. The Belarus trial did not find an association (4). We conclude that breastfeeding may provide some protection against overweight or obesity, but residual confounding cannot be ruled out.

Intelligence tests. Breastfeeding was associated with an increase in 3.5 points in normalized test scores in the pooled analyses of all studies, and 2.2 points when only the high-quality studies are included. The two randomized trials also found significant effects (1,5). We conclude that there is strong evidence of a causal effect of breastfeeding on IQ, although the magnitude of this effect seems to be modest.

References

- 1 Lucas A, Morley R, Cole TJ, et al. Breast milk and subsequent intelligence quotient in children born preterm. *Lancet* 1992; 339: 261–4.
- 2 Lucas A, Morley R. Does early nutrition in infants born before term programme later blood pressure? *BMJ* 1994; 309: 304–8.
- 3 Singhal A, Cole TJ, Fewtrell M, et al. Breastmilk feeding and lipoprotein profile in adolescents born preterm: follow-up of a prospective randomised study. *Lancet* 2004; 363: 1571–8.
- 4 Kramer MS, Matush L, Vanilovich I, et al. Effects of prolonged and exclusive breastfeeding on child height, weight, adiposity, and blood pressure at age 6.5 y: evidence from a large randomized trial. *Am J Clin Nutr* 2007; 86: 1717–21.
- 5 Kramer MS, Aboud F, Mironova E, et al. Breastfeeding and child cognitive development: new evidence from a large randomized trial. *Arch Gen Psychiatry* 2008; 65: 578–84.

For further information please contact:

Department of Maternal Newborn Child and Adolescent Health (MCA)

World Health Organization

20 Avenue Appia, 1211 Geneva 27, Switzerland

Fax: +41 22 791 48 53,

E-mail: mncah@who.int

Web site: http://www.who.int/maternal_child_adolescent



ISBN 978 92 4 150530 7



RESEARCH ARTICLE

Open Access

Prevalence and outcomes of breast milk expressing in women with healthy term infants: a systematic review

Helene M Johns^{1,2*}, Della A Forster^{1,2}, Lisa H Amir¹ and Helen L McLachlan^{1,3}

Abstract

Background: Expressing breast milk has become increasingly prevalent, particularly in some developed countries. Concurrently, breast pumps have evolved to be more sophisticated and aesthetically appealing, adapted for domestic use, and have become more readily available. In the past, expressed breast milk feeding was predominantly for those infants who were premature, small or unwell; however it has become increasingly common for healthy term infants. The aim of this paper is to systematically explore the literature related to breast milk expressing by women who have healthy term infants, including the prevalence of breast milk expressing, reported reasons for, methods of, and outcomes related to, expressing.

Methods: Databases (Medline, CINAHL, JSTOR, ProQuest Central, PsycINFO, PubMed and the Cochrane library) were searched using the keywords milk expression, breast milk expression, breast milk pumping, prevalence, outcomes, statistics and data, with no limit on year of publication. Reference lists of identified papers were also examined. A hand-search was conducted at the Australian Breastfeeding Association Lactation Resource Centre. Only English language papers were included. All papers about expressing breast milk for healthy term infants were considered for inclusion, with a focus on the prevalence, methods, reasons for and outcomes of breast milk expression.

Results: A total of twenty two papers were relevant to breast milk expression, but only seven papers reported the prevalence and/or outcomes of expressing amongst mothers of well term infants; all of the identified papers were published between 1999 and 2012. Many were descriptive rather than analytical and some were commentaries which included calls for more research, more dialogue and clearer definitions of breastfeeding. While some studies found an association between expressing and the success and duration of breastfeeding, others found the opposite. In some cases these inconsistencies were compounded by imprecise definitions of breastfeeding and breast milk feeding.

Conclusions: There is limited evidence about the prevalence and outcomes of expressing breast milk amongst mothers of healthy term infants. The practice of expressing breast milk has increased along with the commercial availability of a range of infant feeding equipment. The reasons for expressing have become more complex while the outcomes, when they have been examined, are contradictory.

* Correspondence: h.johns@students.latrobe.edu.au

¹Mother & Child Health Research, La Trobe University, Melbourne, Victoria, Australia

²Royal Women's Hospital, Parkville, Victoria, Australia

Full list of author information is available at the end of the article

Background

Although data are collected about the proportion of women breastfeeding on discharge from hospital, little is known about how many women are expressing to provide breast milk feeds in addition to, or as an alternative, to feeding directly at the breast. There has been some discussion about increasing numbers of women in Australia, United States of America, the United Kingdom and Singapore expressing to give breast milk feeds rather than breastfeeding directly from the breast [1-6]. Only two studies, one conducted in Australia and one in Singapore [2,6], measured expressing over time. Both reported an increase [2,6].

From an historical point of view, Fildes' 1986 publication about the history of infant feeding provides a comprehensive insight into infant feeding practices from antiquity and describes related medical practices, popular customs and beliefs [7]. The 'drawing off' of breast milk was discussed by Avicenna (AD 980-1036) in the context of milk that was believed to be unpleasant smelling or too thick for the baby to drink [7]. Subsequent references to expressing describe the sucking glass, first mentioned in the mid-16th century [7,8]. The mother applied a glass cup to her breast and sucked on the end of its long glass stem to express milk when her nipples were cracked, or her breast inflamed or infected. During a time when there was concern about the undesirable effects of feeding colostrum to the newborn in pre-industrial Europe, the sucking glass was used as an alternative to employing children or puppies to remove this early milk while the baby was fed by a wet nurse [7].

Developments in breast pump design and uptake over the last century are reflected in changes in 'brand' or company names during the same period. A collection at the Powerhouse Museum in Sydney, Australia includes the *Breast Exhauster* (1892), the *Breast Reliever* (1947), the more recent *Kaneson* hand pump (1973) [9] and the water operated *Ellis Expressor* (1970), locally designed to be connected to a kitchen tap [10]. By the early 1980s, breast pumps were transformed, as the red rubber tubing and glass apparatus and, in the case of the electric breast pump, the noisy motor, were substituted for more appealing designs. Pastel colours, discreet motors and less angular shapes became the norm and these pumps are now promoted with names that are arguably designed to enhance market acceptability. In addition to those mentioned in the previous paragraph, examples in name and design are seen in the earlier *Lopuco* and *Egnell electric* [11] and their successors, the *Diana*, *Free-style*, *Pump-in-style*, *Symphony*, *Swing*, *Harmony*, *Elite* and *Purely Yours* pumps [12,13].

In the world of parenting print media, breast pumps have a growing advertising presence. A hand search of the catalogue of *Essence*, the bi-monthly member

magazine of the Australian Breastfeeding Association (ABA) demonstrates a change in the focus of consumer discussion about breastfeeding over time. Breastfeeding is convenient, and advertising for breast pumps may be interpreted as suggesting that expressing is equally so. Blum writes of "the new regularized, fetishized breastfeeding . . . exemplified in the widespread advertising of pumps" [14] (p. 55). Breast milk expression appears to have become more popular as the associated equipment has become more sophisticated and readily available.

Many of the studies about expressing breast milk focus on premature and/or unwell infants [15-17] reflecting the main reasons women expressed to feed their infants in the past. It is likely that up until the last 20 years healthy term infants were either breastfed or bottle fed with infant formula. Although more recent literature has discussed the prevalence of breast milk expression and suggested that more women are expressing their milk [1,5], measurement of this phenomenon is limited and the consequences relatively unknown.

Defining breastfeeding is complex. Discussion has previously focused on the accurate measurement of breast milk feeding; its exclusivity and duration [18]. That is, breastfeeding was the term used to describe any breast milk intake regardless of the mode of its delivery. The focus of recent debate has shifted and the emerging popularity of expressing presents another complexity; the need to find out how breast milk is given, directly at the breast, or otherwise [19]. In addition, Geraghty and Rasmussen have recommended a need to identify at what age the infant is exposed to expressed breast milk, and whose milk is being used [20].

In this paper *expressing* (also known as *pumping*) is used to describe using a pump to obtain breast milk, and *hand expressing* is used for instances where expressing is done by hand. *Breastfeeding* is used to describe the act of feeding directly from the breast, and *breast milk feeding* includes any means by which breast milk is given to the infant.

The aim of this paper is to systematically explore the literature related to breast milk expressing by women who have healthy term infants, including the prevalence of breast milk expressing, and the reported reasons for, methods of, and outcomes related to expressing.

Methods

The literature search for this paper included a search of Medline, CINAHL, JSTOR, ProQuest Central, PsycINFO, PubMed databases and the Cochrane library with no limit on the year of publication. Reference lists of identified papers were also examined. A hand search of consumer-focused breastfeeding newsletters was undertaken at the ABA Lactation Resource Centre in Melbourne, Australia which holds a collection of

more than 18,000 documents related to human lactation. Relevant media and conference proceedings and personal communications were also searched. Only English language papers were included. Keywords used were: milk expression, breast milk expression, breast milk pumping, prevalence, outcomes, statistics and data. The date of the most recent electronic search was 26 February 2013.

Most of the articles identified in the search were specifically about expressing breast milk for sick and/or premature infants and therefore not relevant for this review, however these groups were included in the original search to ensure all relevant articles were located. A flowchart was developed according to PRISMA guidelines to summarise articles obtained in the literature search [21]. This tool is used to illustrate how many references have been located, the number of exclusions and the criteria for and number of eventual inclusions in the completed review.

All papers about expressing breast milk for healthy term infants were considered for inclusion. Papers about expressing that focused exclusively on premature infants were not included except where relevant for other aspects of this literature review (eg, Methods of expressing).

In the Results, the literature has been classified according to: prevalence of breast milk expressing, methods of expressing and reasons why women express. In addition outcomes and other implications of expressing are included. In each section the type and quality of papers identified is described and the papers are summarised and presented in tables under topic headings.

Results

A total of 22 papers identified were relevant to breast milk expression, but only seven papers reported the prevalence and/or outcomes of expressing amongst mothers of well term infants. Figure 1 provides a visual representation of the publications identified and reviewed [21]. All of the included papers were published subsequent to 1999. Many papers were descriptive rather than analytical and some were commentaries [22-24], including calls for more research, more dialogue and clearer definitions of infant feeding practice [20,22,25,26].

Included papers are summarised in tables, in chronological order, under the sub-headings:

- Prevalence of breast milk expressing;
- Methods of expressing;
- Reasons why women express their milk;
- Impact of expressing on breastfeeding outcomes;
- Other implications of expressing.

Prevalence of breast milk expressing

Seven papers were identified that reported on the prevalence of expressing amongst mothers of well, term infants. They came mainly from USA and Australia [1-6,27]. The papers vary in quality and design (Table 1).

Although several papers refer to an *increase* in the prevalence of expressing, the only data that actually documented such an increase were from Western Australia and Singapore [2,6]. The Perth Infant Feeding Study I (PIFS I), conducted in 1992-93 was followed by the Perth Infant Feeding Study II (PIFS II) ten years later [2]. Binns et al. reported the proportion of mothers

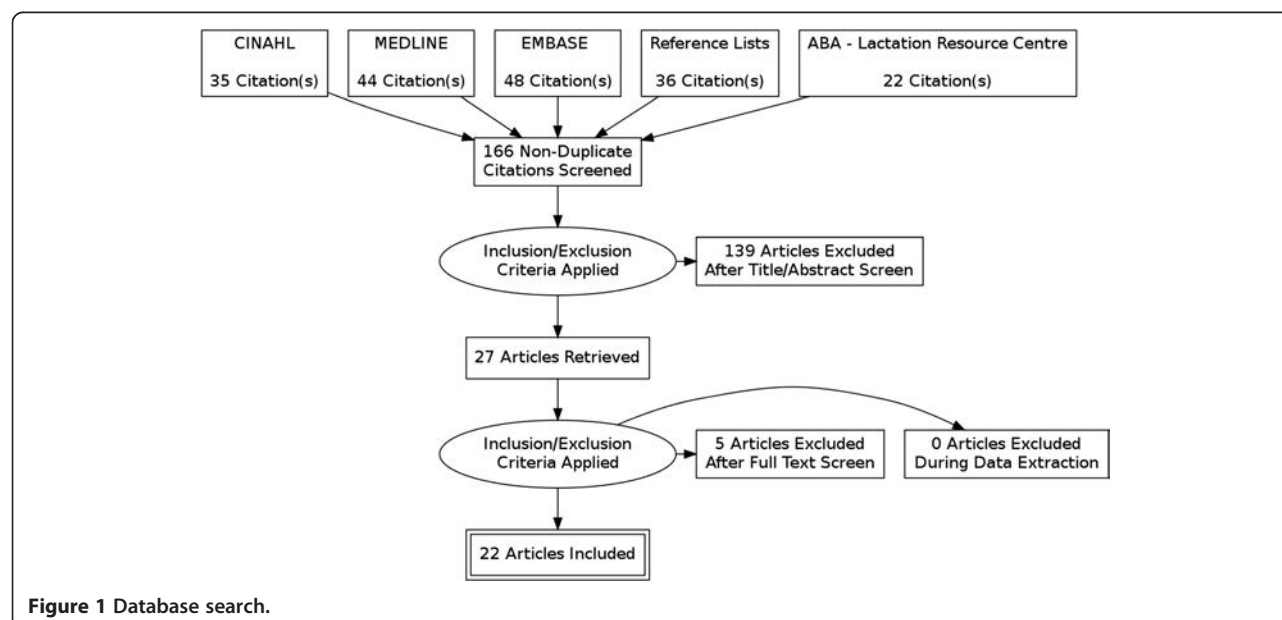


Table 1 Studies exploring prevalence of breast milk expressing

Author, year, country	Design	Location, participants, year of study and recruitment	Study aims and outcome measures	Results	Strengths/Limitations
Geraghty et al. 2005 USA [1]	Retrospective cohort	Cincinnati, Ohio n = 346 2002 Random selection postal recruitment when infants were between 2 and 3 years old	Quantify breast pump use	77% (182/236) ever used a breast pump	Large sample size
			Identify relationships between breast pump use and	No significant difference in expressing between 4 groups of mothers; those of	Clear differentiation between breastfeeding and breast milk feeding
			- singleton vs. multiple pregnancy	- term singletons	Periodic reports re. proportion of expressing versus breastfeeding (at 1 day, 3 days, 2 weeks and monthly until 6 months)
			- gestation at birth	- preterm singletons	<i>Retrospective data, possible recall bias, initial contact made when children between 2–3 years of age-requesting detailed information about infant feeding at 24 hrs, 3 days, 2 weeks and then at monthly intervals to 6 months.</i>
Binns et al. 2006 Australia [2]	Longitudinal	Perth, Western Australia PIFS I n = 556 1992–93 PIFS II n = 587 2002–03 Recruited in hospital in early postpartum period	- Explore determinants of breastfeeding	- term multiples - preterm multiples - 68% (236/346) received at least some breast milk. - 5% (12/236) exclusively expressed to feed – all mothers of premature babies	<i>Expressing methods not distinguished between hand, manual or electric pump</i>
			- Measure and compare prevalence of expressing	- PIFS I 38% (211/556) expressing by 6 weeks - PIFS II 69% (405/587) expressing by 4 weeks	Large sample
				- Expressing rates steadily decline after 6 weeks:	Comparison of similar groups 10 years apart
				26% (145/556) at 24 weeks in PIFS I 28% (164/587) at 22 weeks in PIFS II	<i>Limited detail about expressed breast milk/ breastfeeding proportions.</i>
					<i>Public patients only, perhaps not representative</i>
					<i>Differing interview schedules – less clarity for comparison</i>
Labiner-Wolfe et al. 2008 USA [3]	Longitudinal	National study n = 4606	- Reasons why women express	- Most common reason: for someone else to feed baby	Large sample
			- Amount and prevalence of milk expression	- 85% (1329 /1564) between 1.5 and 4.5 months postpartum had expressed at some time since birth	Maternal recall previous 7 days

Table 1 Studies exploring prevalence of breast milk expressing (Continued)

Shealy et al. 2008 USA [4]	Longitudinal	2005–2007	- Associated socio-demographic factors	- 68% (1015/1493) of this group had expressed in 2 weeks before survey, 25% (373/1493) regularly	Measures frequency of expressing (asked how many times expressed in previous 2 weeks and if expressed on a regular schedule)
		from IFPS II		- Expressing associated with: maternal employment, higher income, first breastfeeding experience	<i>Not nationally representative: Older, more educated, more likely to be white, employed, higher income, less likely to smoke. More likely to breastfeed and for longer</i>
		Mail survey 2, 5 and 7 months postpartum			
		National study	Describe breastfeeding in first 12 months to identify:	- 0.06% of babies fed expressed breast milk exclusively – 2/3 of these ceased breast milk feeding by 4 weeks	Detailed analysis of feeding type/frequency/duration of individual feeds
		n = 2587	- Prevalence of exclusive pumping and formula supplementing		<i>Limited data re. expressing – except when it was exclusive</i>
Clemons & Amir 2010 Australia [5]	Cross-sectional	2005–2007	- Patterns and trends in breastfeeding related to common advice given		
		from IFPS II			
		Monthly postal questionnaires			
		State-wide study, Victoria	- Prevalence of breast milk expression	- 67% (602/898) had fully breastfed prior to hospital discharge	Large study
		n = 903	- Demographic characteristics of women who express, why and how they do it	- 14% (125/898) had breastfed and expressed to feed their baby	<i>Possible selection bias (members of the Australian Breastfeeding Association)</i>
Hornbeak et al. 2010 Singapore [6]	Retrospective cohort	2008	- Women's experience of using breast pumps	- Of those whose youngest child was over six months 95% (628/661) fully breastfed for at least six months	<i>Timing of questionnaire, possible recall bias</i>
		Online questionnaire sent to Australian Breastfeeding Association members who had an email address		- 4% (34/898) expressed and exclusively fed EBM	
		Baby any age		- 98% (885/903) ever expressed	
		Singapore	- Prevalence and patterns of breastfeeding in Singaporean Chinese mothers from birth to 6 months	- Initiation of breast milk feeding increased from 69% (144/210) in 2000-2001 to 82% (538/656) in 2006/2008	Large representative sample of Chinese Singaporean mothers
		n = 3009		- Exclusive expressing increased from 9% (18/144) in 2000-2001 to 18% (118/538) in 2006/2008	<i>Limited detail about expressed breast milk/ breastfeeding proportions.</i>
		2006-2008		- Direct breast milk feeding decreased from 34% (72/210) in 2000–2001 to 27% (142/656) in 2006/2008	<i>Possible recall bias - recruitment 6–72 months after birth</i>

Table 1 Studies exploring prevalence of breast milk expressing (Continued)

		Recruited mothers of 6-72 month Chinese Singaporean children through Strabismus, Amblyopia and Refractive Error in Singaporean Children (STARS) Study			Gestational age not indicated
		Mailed invitation			
		Face-to-face interview			
Geraghty et al. 2012	Prospective longitudinal cohort	Cincinnati, Ohio	- Describe who commences expressing early	- 14% (8/59) commenced some expressing in first week	Prospective design assisting recall
		n = 60	- Explore breastfeeding duration in women who express	- By four weeks: 63% (37/59) expressing	Initial weekly visits; used weekly and 24 hour recall to enquire about feeding and expressing
		2004–2007		- Expressing had no effect on duration of breast milk feeding	Clear differentiation between 'breastfeeding' and 'breast milk feeding'
		Recruited face to face in first week after hospital discharge.			Small study
					Recruitment of women who planned to breastfeed for 6 months or more
					Mothers recruited for study knew they were going to be assisted to pump and may have been more likely to be comfortable with this.
					Possible introduction of bias as weekly collection of breast milk was initiated at 1 week by research nurse using an electric breast pump

who expressed breast milk (not necessarily expressing exclusively) during the first six weeks after birth, and found a 31% increase over ten years as well as a gradual decline in expressing after the first six weeks [2]. (The Singapore study reports exclusive expressing and is discussed below [6]).

In Mid-Western USA in 2002, Geraghty et al. found only sixteen percent (55/346) of women breastfed exclusively at the breast for the duration of their lactation and just seven percent (24/346) of the same group breastfed exclusively at the breast for a minimum of six months [1]. The authors concluded that expressing had become an integral part of human milk feeding [1]. Also in the USA, Labiner-Wolfe et al. analysed data from the Infant Feeding Practices Study II (IFPS II) (collected between 2005 and 2007), and found that 85% (1329/1564) of breastfeeding women had expressed breast milk, and that of these women, more than half had done so in the first week after birth [3]. An online study in Victoria, Australia found that 98% (885/903) of breastfeeding women had expressed at some time [5], however, this was a very select group – as respondents were all members of the ABA.

A small proportion of women never breastfed, but rather exclusively breast milk feed using expressed breast milk. Geraghty et al. reported that five percent of women exclusively expressed and fed breast milk, all of whom were mothers of premature infants [1]. Another study in the USA identified a six percent exclusive expressing rate amongst infants between 35 and 45 weeks gestation [4]. In Australia, Clemons and Amir found that four percent of women in their cross-sectional study expressed exclusively [5]. In comparison, the exclusive expressing group was higher in Singapore; increasing from nine to eighteen percent between 2000–01 and 2006–08, apparently at the expense of direct breastfeeding which decreased from 34 to 22 percent over the same time period [6].

Methods of expressing

A Cochrane review, which included 23 studies assessing breast milk expression methods found that there was no difference between manual and electric pumps in terms of breast milk production [28]. Most of the studies in the systematic review were excluded from this review because they did not meet inclusion criteria, mainly because they focused on premature or unwell infants. The papers discussed in this section are described in Table 2.

Given its universal accessibility, the simplest way to express milk is by hand, and evidence from a Japanese study demonstrates that this is the most effective method to use when expressing colostrum [29]. Hand expressing however, was associated with increased reports of local pain compared to electric breast pumping [29]. A randomised trial in the USA which compared

hand expressing and pump use found that early hand expression appeared to improve breastfeeding rates at two months when compared with using a pump [30].

Although hand expression may be just as effective [28], and electric breast pumps are considerably more expensive than simple hand expression or the use of a hand operated pump, the use of electric breast pumps has become more popular over the last decade [31], and in Perth, Australia, has increased threefold in ten years [2]. Electric breast pumps are a regular feature of post-natal wards of maternity hospitals in Washington, DC where Buckley, examining the views of lactation consultants about breast pumps, writes about the prominence of the breast pump, evidenced by the universality of its provision, in her words, “*A breast pump for every room*” [31] (p.16). The accessibility of electric breast pumps is demonstrated in an online study in Australia which found that 66% (556/843) of breastfeeding women had used an electric breast pump [5]. In a mail survey of 3,606 women from the IFPS II in the USA, women who undertook regular scheduled expressing were more likely to use electric breast pumps [3].

Reasons why women express their milk

Factors that appear to be associated with women's decisions to express their breast milk have been reported in some papers (Table 3). Women who experience difficulty establishing breastfeeding are more likely to express [2,3,5,27], and mothers with premature or low birth weight infants, mothers who are unwell, those who have not previously breastfed are also more likely to express [5,27].

Women with an elevated body mass index (BMI) are more likely to express their milk than to breastfeed, perhaps related to anxiety about exposing their bodies [27,32]. Obese women often have large breasts and may experience difficulty feeding discretely [33]. In addition these women may express because of physical difficulty with breastfeeding. Large breast size may impede maternal ability to see and or facilitate appropriate infant attachment and feeding [33]. Leonard et al. investigated breast milk expressing behaviours and concluded that expressing may support longer durations of breastfeeding in overweight or obese women [32]. Embarrassment about breastfeeding in public has been identified as a reason women express regardless of cultural background or body size [5,34,35]. Cultural differences may inhibit women from breastfeeding outside the home, leading to some women expressing so that they can avoid exposing their bodies in public [3,5,34].

Other reasons women express include breastfeeding problems such as mastitis and breast engorgement [2]; nipple pain and difficulty with attachment to the breast [5]; concern about oversupply or undersupply [3,5,34,36]

Table 2 Studies exploring methods of expressing

Author, year, country	Design	Location, participants, year of study and recruitment	Study aims and outcome measures	Results	Strengths/Limitations
Binns et al. 2006 Australia [2]	Longitudinal	Perth, Western Australia	- Explore determinants of breastfeeding	- Approx. 60% (n = 1143) using manual pumps in both studies	Large study
		PIFS I n = 556	- Measure and compare prevalence of expressing 1992-93 and 2002-03	- Use of electric pumps increased by 31% in 10 years	Comparison of similar groups 10 years apart
		1992-93			<i>Actual figures not given</i>
		PIFS II n = 587			<i>Public patients only, perhaps not representative</i>
Labiner-Wolfe et al. 2008 USA [3]	Longitudinal	2002-03			
		Recruited in hospital in early post-partum period			
		National study n = 4606	- Reasons why women express - Amount and prevalence of milk expression	- Birth to 1.5 -4.5 months 80% (105/1302) battery or electric 44% (573/1302) manual pump 14% (18/1302) hand	Large sample size 3 mailed questionnaires seeking information re. feeding in previous fortnight - recall bias unlikely
		2005-2007 IFPS II Mail survey 2, 5 and 7 months postpartum	- Associated socio-demographic factors	- Previous 3 months to 6.5-9.5 months 73% (39/529) battery or electric) 33% (18/529) manual pump 13% (69/529) hand	Detailed information re. methods of expression over time <i>Not nationally representative, older, more educated, more likely to be white, employed, higher income, less likely to smoke. More likely to breastfeed and for longer</i>
Ohshima et al 2010 Japan [31]	Sequential crossover	Yokohama, Kanagawa	- Comparison of effectiveness and comfort of manual and electric breast expression in first 48 hours after birth	- Manual expressing associated with greater milk volume: net milk yield per woman 2 ml.	Limited other exploration of this area
		n = 11		- Manual expression 2 ml (median; range: 0-12.6 ml.)	<i>Small study</i>
		2003-2004		- Electric expression 0.6 ml. (0-7.2 ml.) (P < 0.05).	<i>Infant gestation and health status not indicated</i>
		Mothers of infants admitted to neonatal intensive care recruited in hospital soon after birth		- Manual pump associated with more reports of pain	
Flaherman et al 2012 USA [32]	RCT	San Francisco & Sacramento, California	Comparison of hand and electric expression measured;	- At 2 months mothers assigned to hand expressing were more likely to be breastfeeding (97%, 47/48) than mothers assigned to breast pumping (73%,35/48) (RR:1.32, 95% CI 1.01,1.73)	Limited other exploration of this area, no previous studies linking type of expressing to breastfeeding outcomes
		n = 68	- Milk transfer		Thorough discussion
		2007-2009	- Breast pain		<i>Small study, final outcome assessment based on 48 participants</i>
		Recruited 12-36 hours after birth	- Breastfeeding confidence		

Table 2 Studies exploring methods of expressing (Continued)

					Possible bias- recruited infants experiencing feeding difficulty
			- Breast milk expression experience		No control group
			- Breastfeeding rates at 2 months		
Becker et al. 2011 UK [30]	Systematic review	International	- Assessment and review of randomised and quasi randomised trials comparing methods of milk expression any time after birth and crossover trials commencing at least 28 days after birth	- More milk with relaxation tape	Systematic review
		n = 642 women from 23 studies		- No difference in mean vol. with simultaneous or sequential pumping, or between manual and electric pumps studied	Most studies specifically related to the care of the pre-term infant

Table 3 Studies exploring reasons women express

Author, year, country	Design	Location, participants, year of study and recruitment	Study aims and outcome measures	Results	Strengths/Limitations
Dykes & Williams 1999 UK [34]	Longitudinal, phenomenological study	Northern England, n = 10 1998 Postnatal primiparas recruited face-to-face in hospital, home visits at 6, 8 & 12 weeks	- Explore women's experience of expressing particularly perception of adequacy of milk supply	- Beliefs re. adequacy of breast milk supply influenced by interplay of feeding management, infant behaviour, lactation physiology and maternal mental health.	<i>Small mono-cultural group</i>
Binns et al. 2006 Australia [2]	Longitudinal cohort	Perth, Western Australia PIFS I n = 556 1992–93 PIFS II n = 587 2002–03 Recruited in hospital in early post-partum period.	- Explore determinants of breastfeeding - Measure and compare prevalence in expressing	- Early breastfeeding difficulties, - Engorgement, sore nipples, mastitis - Feed to be given by someone else - To store extra milk - Father to feed - To increase supply - Feeding/attachment problems - To get baby to drink from a bottle - Just to try it out	Comparison of similar groups 10 years apart <i>Mainly women who expressed to manage breastfeeding difficulties</i> <i>Public patients only, perhaps not representative</i>
Labiner-Wolfe et al. 2008 USA [3]	Longitudinal cohort	National study n = 3606 2005–2007 from IFPS II	- Reasons why women express - Amount and prevalence of milk expression - Associated socio-demographic factors	- to allow someone else to feed - maternal employment - to have an emergency milk supply - no previous breastfeeding experience - geographic location (Midwest Vs. West) - embarrassed to breastfeed in public	Large sample <i>Not nationally representative</i> <i>Participants older, more likely to be educated, white, employed, higher income</i>
Buckley 2009 USA [33]	Focus groups	Washington, DC n = 12	- Ascertain lactation consultant's beliefs and experiences re. impact of breast pumps on breastfeeding practice	- Technological birth contributes to technological breastfeeding - Engorgement, plugged ducts, to increase supply, to stimulate the let-down reflex, to pull out inverted nipples.	Exploration of professional attitudes to change in feeding practice -no previous exploration of this area <i>Small sample size</i>

Table 3 Studies exploring reasons women express (Continued)

		Lactation consultants		- Return to work	Volunteer participants
		Purposeful sampling		- Measuring milk, diminished confidence in ability to provide enough milk	Date of study not indicated
Clemons & Amir 2010 Australia [5]	Cross-sectional	State-wide study, Victoria	- Prevalence of breast milk expression	- Premature baby/sick mother or baby	Large study
		n = 903	- Demographic characteristics of women who express, why and how they do it	- Attachment problems/not drinking well	Possible selection bias (members of ABA)
		2008	- Women's experience of using breast pumps	- Advised	Timing of questionnaire, possible recall bias
		Online questionnaire sent to Australian Breastfeeding Association (ABA) members who had an email address		- Not enough milk/To store extra milk - Nipple pain - Engorged breasts/mastitis - So someone else can feed baby - Maternal work - Just to try it out - To allow mother to drink alcohol - Uncomfortable breastfeeding in public	
Geraghty et al. 2012 USA [29]	Prospective longitudinal cohort	Cincinnati	- Duration of breast milk feeding	- Planned return to work by 6 months	Prospective design
		n = 60	- Describe who commences expressing early		Small study
		2004–2007			Recruitment of women who planned to breastfeed for 6 months or more
		recruited face to face			Mothers recruited for study knew they were going to be assisted to pump and may have been more likely to be comfortable with this. Possible introduction of bias as weekly collection of breast milk was initiated at 1 week by research nurse using an electric breast pump

and allowing the baby to be fed by someone other than his/her mother [2,3,34,37]. Women express in order to return to paid work [3,5,38]. Those who are in paid employment are more likely to express their milk when there are flexible work arrangements and designated places to express [39,40]. In addition to women who are in paid employment expressing milk, women who have a high income are also more likely to do so [3].

Impact of expressing on breastfeeding outcomes

There have been contradictory reports regarding the association between expressing breast milk and the success and duration of breastfeeding (Table 4). Some studies suggest that expressing to feed (as opposed to breastfeeding solely at the breast) in the early postpartum period is associated with shorter duration of breastfeeding [1,41], while others have found the reverse [42,43]. Binns et al. report on trends in the expression of breastmilk and conclude "The appropriate use of expressed breastmilk allows a mother to achieve six months of exclusive breastfeeding while giving her more options with regards to paid work or study and the management of breastfeeding difficulties"[2] (Page 8). Women who fed at the breast only were found by Schwartz et al. to breastfeed for longer [41] and Chapman et al. reported that expressing did not improve milk volumes or duration of breastfeeding [44]. On the other hand, Win et al. in Perth, Australia, explored the association between expressing and the duration of breastfeeding and reported that mothers who expressed were more likely to be breastfeeding at six months than those who didn't express [45].

Studies that looked at duration of breast milk feeding have had differing outcomes. Ortiz et al. explored the duration of breast milk expression for women allocated to a lactation program which provided equipment and support for expressing. Findings indicated that these women were more likely to breast milk feed for longer than those who did not receive such support [39]. Geraghty et al. found that mothers who fed solely at the breast, particularly in the early weeks postpartum, were more likely to breast milk feed for longer than women who had combined breastfeeding with expressing [1].

Implications for maternal health

Breast pain and nipple trauma have been associated with expressing [5,29]; they contribute to maternal discomfort and distress, and nipple trauma is known to be associated with the development of mastitis [46]. While Thorley identifies compromised mother/infant skin-to-skin contact and bonding as a result of expressing [23], Johnson et al. call breast pumping liberating, giving the mother a means to "... *negotiate some independence and manage the demands of breastfeeding*" [34] (p. 900). These

authors suggest that expressing may facilitate maternal independence, and give the mother freedom from the demands of her baby [34].

Considering the paucity of discussion in the literature, it appears that the implications of expressing on maternal mental health warrants attention.

Other implications of expressing

A range of other possible outcomes of breast milk expression bear consideration and can be seen in Table 5. The additional handling involved in the expression, storage and subsequent bottle feeding of expressed milk creates additional risks for infection in the infant, as discussed and illustrated visually by Geraghty [25]. Breast milk, frozen and fed later loses vitamin content, and, like infant formula, it is at risk of contamination, as it is subject to more handling through the process of preparation [47]. Freezing, defrosting and reheating and microwaving all have the potential to compromise milk quality and safety [48,49].

The infant fed from a bottle, regardless of the type of milk, is deprived of the benefits of self-regulation of intake associated with breastfeeding, which may increase the risk of subsequent childhood obesity [50,51]. Orofacial implications include the risk of dental caries associated with the use of a teat [52] and orthodontic problems associated with not breastfeeding [53].

Discussion

Although there is some commentary about an increase in breast milk expressing in the literature, actual measurement of the phenomena is quite limited. In addition, expressing breast milk is anecdotally less common in countries where there are more generous maternity leave provisions in terms of length of financial support, but there is limited evidence in the literature to support this. Two studies, one in Australia and the other in Singapore provide the only data actually documenting an increase in expressing over time [2,6]. Other authors discuss the prevalence of expressing and describe a preponderance of the practice [1-6,27]. This suggests that breast milk feeding solely at the breast is actually quite rare, at least in the developed world. Some breast milk expressing deserves consideration as incidental, something a mother might do only on occasion, for example when she needs to go out without her baby, when introducing infant cereal, when the infant is ill or unable to attach or is refusing the breast. We know that expressing breast milk has become more common, clarification of the amount, the proportion and the frequency of expressing and breast milk feeding is necessary before we can properly explore the implications of this relatively unexplained shift away from the breast.

Table 4 Impact of expressing on breastfeeding outcomes

Author, year, country	Design	Location, participants, year of study and recruitment	Study aims and outcome measures	Results	Strengths/Limitations
Chapman et al. 2001 USA [45]	RCT	Hartford, Connecticut	Effects of expressing before the onset of lactation :	- No significant difference in milk transfer or breastfeeding duration between women who expressed breast milk and those who did not.	<i>Only women who had a caesarean section</i>
		n = 60	- on early milk transfer	- Primiparous women in pumping group breastfed for 5 months less than those in control group but this finding was not statistically significant.	<i>Study underpowered for primiparous women</i>
		1997–1998 Convenience sample 8–24 hours post Caesarean Section	- on subsequent breastfeeding duration		
Schwartz et al. 2002 USA [42]	Prospective cohort	Detroit, Ann Arbor and Southfield, Michigan and Omaha, Nebraska	- Determine demographic, behavioural and clinical factors associated with weaning from breast in the first 12 weeks	- Michigan women (n = 711) who expressed breast milk were 3 times more likely to wean than those who didn't (Hazard Ratio: 3.0 95% CI 1.3,6.7)	Large study
		n = 946		- Nebraska women (n = 235) showed no association between pumping and weaning (HR: 0.6, 95% CI 0.3,1.5)	<i>Only measured to 12 weeks</i>
		1994–1998			<i>Non-representative sample</i>
		Recruitment:			<i>Michigan group were recruited from an alternative birthing centre and were significantly more likely to be older than 30 years, have a bachelor's degree, have 3 or more children and have had a vaginal birth</i>
		Michigan - at birth centre orientation Nebraska - on maternity leave application to large company			
Ortiz 2004 USA [37]	Clinical audit	Burbank, California	- Duration of breast milk feeding related to a range of employee chosen lactation support options	- 98% (452/ 462) breastfeeding initiation	Large study over 4.5 years
		n = 462		- 74% (246/332) expressed milk until infant at least 6 months	<i>Limited differentiation between breastfeeding and expressing / breast milk feeding</i>
		1993–1999		- 24% (81/332) expressed milk until infant at least 12 months	<i>No information re any other infant feeding/ exclusivity of breast milk feeding</i>
Geraghty et al. 2005 USA [1]	Retrospective cohort	Cincinnati, Ohio	Measure breast pump use	Of breast milk feeding mothers:	Large sample size
		n = 346	Identify relationships between breast pump use and:	- 10% (24/346) breastfed exclusively for a minimum of 6 months	Breastfeeding / breast milk feeding clearly differentiated
		2002	- singleton vs. multiple pregnancy		

Table 4 Impact of expressing on breastfeeding outcomes (Continued)

		Random selection Postal recruitment when infants were between 2 and 3 years old	- gestation at birth - breastfeeding outcomes	- 16% (55/346) breastfed exclusively for duration of their breast milk feeding - 77% (182/236) expressed at some time in first 6 months - 59% (140/236) ceased breast milk feeding by 6 months Of the 140 women who had ceased breast milk feeding by 6 months, at the time point just prior to exclusive formula feeding: - 76% (106/140) were either expressing exclusively or combining expressing with breastfeeding - 24% (34/140) were breastfeeding Early breastfeeding associated with a longer duration of breast milk feeding	Periodic reports re. proportion of expressing versus breastfeeding (at 1 day, 3 days, 2 weeks and monthly until 6 months) <i>Retrospective data, possible recall bias as participants were recruited 2 or more years post birth</i>
Win et al. 2006 Australia [55]	Prospective cohort	Perth, Western Australia PIFS II n = 587 2002–03 Recruited in hospital at birth.	- Investigate association between breast milk expression and breastfeeding duration	- Mothers who expressed at least once more likely to be breastfeeding at 6 months (RR: 0.71, 95% CI 0.52,0.98)	Prospective design assisting recall <i>Ever “expressed” / “any” breastfeeding</i> <i>? lower socio economic bias</i> <i>No account of frequency of expressing</i>
Meehan et al. 2008 USA [43]	Quasi-experimental	Los Angeles, California n = 208 2001 Breast pump loan program for low income Women with Children (WIC) recipients	- Evaluation of program to facilitate breastfeeding for low income mothers - Maternal request for formula from WIC program used as proxy measurement to give indication of partial breastfeeding	- Electric pump loan associated with more breastfeeding at 6 months. Mothers loaned a breast pump 5.5 times more likely to than those who hadn't received one to not have requested formula by 6 months (OR: 5.5, 95% CI 2.0,15.1)	<i>Limited reliability of proxy measurement to assess breast milk feeding prevalence or duration</i> <i>No differentiation between breastfeeding and breast milk feeding</i>
Fein 2008 USA [41]	Prospective cohort	National n = 810	- Examine strategies used to combine work and breastfeeding	Median duration of breast milk feeding associated with workplace practices:	Large National study Prospective design

Table 4 Impact of expressing on breastfeeding outcomes (Continued)

			- Identify strategies associated with enhanced breastfeeding intensity/longer duration		
		2005–2007		- expressing and breastfeeding (32.4 weeks) (n = 75)	Questionnaire design with 7 day recall
		from IFPS II		- breastfeed at the breast only (31.4 weeks) (n = 250)	No description of feeding method away from workplace
		Recruitment via postal questionnaire in late pregnancy		- expressing only (26.3 weeks) (n = 75)	Older, less educated, low income and women from racial/ethnic minority groups underrepresented
				- neither breastfeeding or expressing (14.3 weeks) (n = 128)	
Clemons & Amir 2010 Australia [5]	Cross-sectional	State-wide, Victoria	- Prevalence of breast milk expression	- 27% (218/903) indicated that expressing had allowed them to breastfeed for longer	Large study
		n = 903	- Demographic characteristics of women who express, why and how they do it		Possible selection bias (members of ABA)
		2008	- Women's experience of using breast pumps		Timing of questionnaire, possible recall bias
		online questionnaire			
		ABA members with internet addresses			
Dabritz et al. 2010 USA [56]	Retrospective cohort	Yolo County, California	- Assess relationship between maternal experience in hospital and any breastfeeding at six months	- Almost exclusive breastfeeding at 6 months associated with not using a breast pump in hospital 77% (93/121) compared to 21% (25/121) who did use a pump in hospital (OR: 0.6 95% CI 0.3,1.0)	Differentiation between breastfeeding and breast milk feeding unclear
		n = 382			
		2006–07			Possible recall bias - interviews 6–9 months after birth
		Recruited in community after birth - 8 months			
Hornbeak et al. 2010 Singapore [6]	Retrospective cohort	Singapore	- Record prevalence and patterns of breastfeeding in Singaporean Chinese mothers	- Breast milk feeding initiation and duration increased over time and were independently associated with higher maternal education, increased milk expression and complementary feeding	Large representative sample of Chinese Singaporean mothers
		n = 3009		Changes between 2000–01 and 2006–08:	Possible recall bias - recruitment 6–72 months after birth
		2006–2008		Infant formula feeding 31% (66/210) to 18% (118/656)	Gestational age not indicated

Table 4 Impact of expressing on breastfeeding outcomes (Continued)

Geraghty et al. 2012 USA [29]	Prospective cohort	Recruited mothers of 6–72 month Chinese Singaporean children through STARS		Breast milk feed initiation 69% (144/210) to 82% (538/656)	
		Mailed invitation		Expressed breast milk 9% (18/210) to 18% (118/656)	
				Combination feeding 26% (54/210) to 41% (269/656)	
		Cincinnati, Ohio	- Determine who expresses their milk by end of 4 weeks and how long they continue feeding	- Milk expression common in first month postpartum	Prospective design
		n = 60		- Milk expression by 4 weeks did not significantly influence duration of breast milk feeding	Clear differentiation between breastfeeding and breast milk feeding
		2004–2007			<i>Recruitment of women who planned to breastfeed for 6 months or more</i>
		Participants enrolled in a research human milk bank recruited at home in first week postpartum			<i>Mothers recruited for study knew they were going to be assisted to pump and may have been more likely to be comfortable with this.</i>
					<i>Possible introduction of bias as weekly collection of breast milk was initiated at 1 week by research nurse using an electric breast pump</i>

Table 5 Other implications of expressing

Author, year, country	Design	Location, participants, year of study and recruitment	Study aims and outcome measures	Results	Strengths/Limitations
Clemons & Amir 2010 Australia [5]	Cross sectional	State-wide study, Victoria n = 903 2008 Online questionnaire sent to Australian Breastfeeding Association (ABA) members who had an email address Baby any age	- Prevalence of breast milk expression - Demographic characteristics of women who express, why and how they do it - Women's experience of using breast pumps	- 17% (126/737) experienced nipple pain associated with pump use	Large study <i>Possible selection bias (members of ABA)</i> <i>Timing of questionnaire, possible recall bias</i>
Li et al. 2010 USA [49]	Longitudinal cohort	National study n = 1597 2005–2007 from IFPS II	- Test infant ability to self-regulate intake – compare active sucking (breastfeeding) with passive feeding (EBM via bottle) - Complete emptying of bottle or cup in late infancy used to indicate self-regulation	- Infants bottle fed early more likely to empty bottle/cup in late infancy - bottle a totally different feeding mode	Large national longitudinal study Minimal reporting bias for exposure and outcome – 7 day retrospective recall Multivariate analysis <i>Maternal report of feeding behaviour/bottle emptying - reporting error possible</i>
Li et al. 2012 [50]	Longitudinal cohort	USA n = 1899 IFPS II 2005–2007	- Multi level analysis to estimate weight gain X type of milk & feeding mode at 3,5, 7, and 12	- Among infants fed only breast milk, Breast milk fed infants gained 780g per month in the first year compared with breastfed infants who gained 729g - Possible association between bottle feeding EBM and increased weight gain	Large national longitudinal study Minimal reporting bias for exposure and outcome – 7 day retrospective recall
Geraghty et al. 2012 USA [28]	Retrospective cohort	Cincinnati, Ohio n = 40 2008 Outpatients attending breastfeeding clinic, recruited by mail	- Examination of methods of maternal expression and infant consumption of breast milk	- All expressed, all babies fed some expressed milk 95% (38/40) infants breastfed and EBM 37% (15/40) fed EBM same day 30% (12/40) fed EBM same week 25% (8/40) fed EBM 1 and 4 weeks later 13% (5/40) fed EBM more than 4 weeks later	Limited other exploration of this area <i>Small size</i> <i>Retrospective data collection</i> <i>Possible recall bias</i>

Although concluding remarks in the 2011 Cochrane review identify the relative effectiveness of hand expressing and less expensive breast pumps [28], it is not surprising that regular scheduled expressing is associated with the use of electric breast pumps [3]. Hand expressing barely rates a mention in the literature about expressing breast milk, although breast pump use and ownership are commonly discussed [2,31,54,55]. Thorley describes breast pumps as a substitute for the skill of hand expression [23]. A recent paper from Flaherman et al. reports on an apparent positive effect of hand expressing when compared with bilateral electric pumping [30]. The authors discuss the possibility that hand expressing contributes to less awkwardness or embarrassment for the mother, who is more likely to be comfortable hand expressing than using a pump when others are present [30]. After the establishment of a mature milk supply however, a 2013 review of studies comparing different methods of milk expression [37], identifies several that found electric breast pumping to be more effective than any other method in terms of milk volume obtained [49,56,57]. The let-down reflex, a physiological response that is the process by which milk becomes available for the baby, can be inhibited by stressful situations such as embarrassment [58]. This may also be the case when unfamiliar equipment such as the breast pump is used. Suggestions of unquantifiable benefits associated with teaching hand expressing are made by Morton as she cautions against an “*over-reliance on mechanical interventions*” [59] (p. 276). Maternal confidence may be enhanced by a more relaxed early postpartum experience. Ease with the mechanics of breastfeeding may be fostered for the woman who has had some experience handling her breasts to obtain milk, as she will have done when hand expressing. The confidence attained by the handling and the achievement of actually expressing, as well as the visualisation of her milk may reinforce such confidence as she has expressed simply, with her own hands and without the complication of any additional mechanisation.

Women express breast milk because of doubt about the adequacy of their milk supply [36]. Many do so because of initial difficulties establishing breastfeeding [2,3,5,27]. Maternal return to work has been identified as another reason [3,5,27]. Workplaces which provide options such as on-site childcare, lactation breaks for expressing and/or breastfeeding foster longer term breast milk feeding [60]. Focus on breast expression facilities may however encourage maternal return to the workforce at the cost of other initiatives, or at the cost of broader social change to support women to spend more time with their infants and by implication, breastfeeding. Elevated BMI, cultural differences and embarrassment about breastfeeding in public all contribute to increasing breast milk expression rates [3,5,27,29,32,34].

Simple attribution of cause and effect is probably inadequate in any discussion of expressing and its impact on breastfeeding success and duration. Several authors appear to refer to ‘any’ expressing without specifying relative quantities or proportions of breast milk feeds given, from the breast or otherwise [2,4,6].

Expressing may contribute to a parental focus on the measurement of breast milk. Such quantification of breast milk may undermine confidence about the adequacy of milk supply, and may be reflected in parental anxiety about not knowing how much milk the baby is obtaining when feeding directly from the breast [31].

Conclusions

This literature review has found limited evidence about the prevalence and outcomes of expressing breast milk amongst mothers of healthy term infants. Authors use a variety of definitions to describe the various infant feeding options, which limits our ability to make conclusions. The practice of expressing breast milk has increased along with the commercial availability of a range of infant feeding equipment. Expressing breast milk has become more common and introduces an opportunity for others to feed the baby. It could be argued that, for some families, breastfeeding has been reduced to a task, which is that of providing milk for the baby, quite possibly even beyond arm’s length, something that anyone can do and potentially at the cost to the special relationship between a mother and her infant. The reasons for expressing have broadened and acquired complexity, while the outcomes of expressing, when they have been examined, are contradictory.

Competing interests

The authors declare that they have no competing interests.

Authors’ contributions

HMJ undertook the literature review and wrote the first draft. All authors contributed intellectual input into revisions of the paper. All authors read and approved the final manuscript.

Acknowledgements

HMJ has a La Trobe University PhD scholarship. The authors received no funding for conducting and writing this review.

Author details

¹Mother & Child Health Research, La Trobe University, Melbourne, Victoria, Australia. ²Royal Women’s Hospital, Parkville, Victoria, Australia. ³School of Nursing and Midwifery, La Trobe University, Bundoora, Victoria, Australia.

Received: 5 March 2013 Accepted: 13 November 2013

Published: 19 November 2013

References

- Geraghty SR, Khoury JC, Kalkwarf HJ: Human milk pumping rates of mothers of singletons and mothers of multiples. *J Hum Lact* 2005, **21**(4):413–420.
- Binns CW, Win NN, Zhao Y, Scott JA: Trends in the expression of breastmilk 1993–2003. *Breastfeed Rev* 2006, **14**(3):5–9.

3. Labiner-Wolfe J, Fein SB, Shealy KR, Wang C: **Prevalence of breast milk expression and associated factors.** *Pediatrics* 2008, **122**(2):S63–S68.
4. Shealy KR, Scanlon KS, Labiner-Wolfe J, Fein SB, Grummer-Strawn LM: **Characteristics of breastfeeding practices among US mothers.** *Pediatrics* 2008, **122**(2):S50–S55.
5. Clemons SN, Amir LH: **Breastfeeding women's experience of expressing: A descriptive study.** *J Hum Lact* 2010, **26**(3):258–265.
6. Hornbeak DM, Dirani M, Sham WK, Li J, Young TL, Wong TY, Chong YS, Saw SM: **Emerging trends in breastfeeding practices in Singaporean Chinese women: findings from a population-based study.** *Ann Acad Med Singapore* 2010, **39**:88–94.
7. Fildes V: *Breasts, Bottles and Babies.* Edinburgh: Edinburgh University Press; 1986.
8. Walker M: **Breast pumps and other technologies.** In *Breastfeeding and Human Lactation*. 4th edition. Edited by Riordan J, Wambach K. Sudbury, Massachusetts: Jones and Bartlett; 2010:379–423.
9. Kaneson 100th anniversary webpage. <http://www.kaneson.co.jp/100th/index.html>.
10. Powerhouse Museum Collection Database - Breastpumps. http://www.powerhousemuseum.com/collection/database/search_tags.php?tag=breast+pumps.
11. Green D, Moye L, Schreiner RL, Lemons JA: **The relative efficacy of four methods of human milk expression.** *Early Hum Dev* 1982, **6**(2):153–159.
12. Breast Pump webpage. <http://www.medelabreastfeedingus.com/products/category/breast-pumps>.
13. Ameda webpage - breastfeeding product list. <http://www.breastfeeding.com/breastfeeding-products>.
14. Blum LM: **From Sacred to disembodied motherhood.** In *At The Breast: Ideologies of Breastfeeding and Motherhood in the Contemporary United States*. Boston: Beacon Press; 1999:19–62.
15. Jaeger MC, Lawson M, Filteau S: **The impact of prematurity and neonatal illness on the decision to breast-feed.** *J Adv Nurs* 1997, **25**(4):729–737.
16. Da Silva OP, Knoppert DC, Angelini MM, Forret PA: **Effect of domperidone on milk production in mothers of premature newborns: a randomized, double-blind, placebo-controlled trial.** *CMAJ* 2001, **164**(1):17–21.
17. Fewtrell MS, Loh KL, Blake A, Ridout DA, Hawdon J: **Randomised, double blind trial of oxytocin nasal spray in mothers expressing breast milk for preterm infants.** *Arch Dis Child Fetal Neonatal Ed* 2006, **91**(3):F169–F174.
18. Aarts C, Kylberg E, Hörnell A, Hofvander Y, Gebre-Medhin M, Greiner T, Urmeå U: **How exclusive is exclusive breastfeeding? A comparison of data since birth with current status data.** *Int J Epidemiol* 2000, **29**(6):1041–1046.
19. Hector DJ: **Complexities and subtleties in the measurement and reporting of breastfeeding practices.** *Int Breastfeed J* 2011, **6**:5.
20. Geraghty SR, Rasmussen KM: **Redefining "breastfeeding" initiation and duration in an age of breastmilk pumping.** *Breastfeed Med* 2010, **5**(3):135–137.
21. Moher D, Liberati A, Tetzlaff J, Altman DG, The Prisma Group: **Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement.** *PLoS Med* 2009, **6**(7):e1000097.
22. Rasmussen KM, Geraghty SR: **The quiet revolution: breastfeeding transformed with the use of breast pumps.** *Am J Public Health* 2011, **101**(8):1356–1359.
23. Thorley V: **The dilemma of breastmilk feeding.** *Breastfeed Rev* 2011, **19**(1):5–7.
24. Johnson S, Leeming D, Williamson I, Lyttle S: **Maintaining the 'good maternal body': expressing milk as a way of negotiating the demands and dilemmas of early infant feeding.** *J Adv Nurs* 2012, **69**(3):590–599.
25. Geraghty SR: **Photo album of expressed breastmilk.** *Breastfeed Med* 2010, **6**(6):433–434.
26. Geraghty SR, Sucharew H, Rasmussen KM: **Trends in breastfeeding: it is not only at the breast anymore.** *Matern Child Nutr* 2013, **9**(2):180–187.
27. Geraghty SR, Davidson E, Tabangin M, Morrow A: **Predictors of breastmilk expression by 1 month postpartum and influence on breastmilk feeding duration.** *Breastfeed Med* 2012, **7**(2):112–117.
28. Becker GE, Cooney F, Smith HA: **Methods of milk expression for lactating women.** *Cochrane Database Syst Rev* 2011, **7**:12.
29. Ohyama M, Watabe H, Hayasaka Y: **Manual expression and electric breast pumping in the first 48 h after delivery.** *Pediatr Int* 2010, **52**(1):39–43.
30. Flaherman VJ, Gay B, Scott C, Avins A, Lee KA, Newman TB: **Randomised trial comparing hand expression with breast pumping for mothers of term newborns feeding poorly.** *Arch Dis Child-Fetal Neonatal Ed* 2012, **97**(1):F18–F23.
31. Buckley K: **A double-edged sword: lactation consultants' perceptions of the impact of breast pumps on the practice of breastfeeding.** *J Perinat Educ* 2009, **18**(2):13–22.
32. Leonard SA, Labiner-Wolfe J, Geraghty SR, Rasmussen KM: **Associations between high prepregnancy body mass index, breast-milk expression, and breast-milk production and feeding.** *Am J Clin Nutr* 2011, **93**(3):556–563.
33. Amir LH, Donath SM: **A systematic review of maternal obesity and breastfeeding intention, initiation and duration.** *BMC Pregnancy Childbirth* 2007, **7**:9.
34. Johnson S, Williamson I, Lyttle S, Leeming D: **Expressing yourself: A feminist analysis of talk around expressing breast milk.** *Soc Sci Med* 2009, **69**(6):900–907.
35. Bonuck KA, Trombley M, Freeman K, McKee D: **Randomized, controlled trial of a prenatal and postnatal lactation consultant intervention on duration and intensity of breastfeeding up to 12 months.** *Pediatrics* 2005, **116**(6):1413–1426.
36. Dykes F, Williams C: **Falling by the wayside: a phenomenological exploration of perceived breast-milk inadequacy in lactating women.** *Midwifery* 1999, **15**(4):232–246.
37. Flaherman VJ, Lee HC: **"Breastfeeding" by feeding expressed mother's milk.** *Pediatr Clin North Am* 2013, **60**(1):227–246.
38. Fein SB, Mandal B, Roe BE: **Success of strategies for combining employment and breastfeeding.** *Pediatrics* 2008, **122**(2):S56–S62.
39. Ortiz J, McGilligan K, Kelly P: **Duration of breast milk expression among working mothers enrolled in an employer-sponsored lactation program.** *Pediatr Nurs* 2004, **30**(2):111–119.
40. Whaley SE, Meehan K, Lange L, Slusser W, Jenks E: **Predictors of Breastfeeding Duration for Employees of the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC).** *J Am Diet Assoc* 2002, **102**(9):1290–1293.
41. Schwartz K, Arcy HJS D, Gillespie B, Bobo J, Longeway M, Foxman B: **Factors associated with weaning in the first 3 months postpartum.** *J Fam Pract* 2002, **51**(5):439–444.
42. Meehan K, Harrison GG, Afifi AA, Nickel N, Jenks E, Ramirez A: **The association between an electric pump loan program and the timing of requests for formula by working mothers in WIC.** *J Hum Lact* 2008, **24**(2):150–158.
43. Bai Y, Wunderlich S, Fly A: **Predicting intentions to continue exclusive breastfeeding for 6 months: a comparison among racial/ethnic groups.** *Matern Child Health J* 2011, **15**(8):1257–1264.
44. Chapman DJ, Young S, Ferris AM, Perez-Escamilla R: **Impact of breast pumping on lactogenesis stage II after cesarean delivery: a randomized clinical trial.** *Pediatrics* 2001, **107**(6):E94.
45. Win N, Binns C, Zhao Y, Scott J, Oddy W: **Breastfeeding duration in mothers who express breast milk: a cohort study.** *Int Breastfeed J* 2006, **1**:28.
46. Amir LH, Garland S, Lumley J: **A case-control study of mastitis: nasal carriage of Staphylococcus aureus.** *BMC Fam Pract* 2006, **7**:57.
47. Boo NY, Nordiah AJ, Alfizah H, Nor-Rohani AH, Lim VKE: **Contamination of breast milk obtained by manual expression and breast pumps in mothers of very low birthweight infants.** *J Hosp Infect* 2001, **49**(4):274–281.
48. Hurst N, Meier P, Riordan J: **Breastfeeding the preterm infant.** In *Breastfeeding and human lactation*. 4th edition. Edited by Riordan J, Wambach K. Sudbury, Massachusetts: Jones and Bartlett; 2010:425–470.
49. Garza C, Johnson CA, Harriet R, Nichols BL: **Effects of methods of collection and storage on nutrients in human milk.** *Early Hum Dev* 1982, **6**(3):295–303.
50. Li R, Fein SB, Grummer-Strawn LM: **Do infants fed from bottles lack self-regulation of milk intake compared with directly breastfed infants?** *Pediatrics* 2010, **125**(6):e1386–e1393.
51. Li R, Magadia J, Fein SB, Grummer-Strawn LM: **Risk of bottle-feeding for rapid weight gain during the first year of life.** *Arch Pediatr Adolesc Med* 2012, **166**(5):431–436.
52. Harris R, Nicoll AD, Adair PM, Pine CM: **Risk factors for dental caries in young children: a systematic review of the literature.** *Community Dent Health* 2004, **21**(1):71–85.
53. Emerich K, Wojtaszek-Slominska A: **Clinical practice. Later orthodontic complications caused by risk factors observed in the early years of life.** *Eur J Pediatr* 2010, **169**(6):651–655.
54. Johns HM, Forster DA, Amir LH, McLachlan HL, Moorhead AM, Ford R, McEgan KM: **In-hospital feeding of term infants - Baseline data of a prospective cohort study to explore the prevalence and outcomes of breast milk expressing and breast pump usage.** *J Paediatr Child Health* 2012, **48**:15.

55. Johns HM, Forster DA, Amir LH, Moorhead AM, McEgan KM, McLachlan HL: **Infant feeding practices in the first 24–48 h of life in healthy term infants.** *Acta Paediatr* 2013, **102**(7):315–320.
56. Slusher T, Slusher IL, Biomdo M, Bode-Thomas F, Curtis BA, Meier P: **Electric breast pump use increases maternal milk volume in African nurseries.** *J Trop Pediatr* 2007, **53**(2):125–130.
57. Slusher TM, Slusher IL, Keating EM, Curtis BA, Smith EA, Orodriyo E, Awori S, Nakakeeto MK: **Comparison of maternal milk (breastmilk) expression methods in an African nursery.** *Breastfeed Med* 2012, **7**(2):107–111.
58. Brodribb W: **Breastfeeding Management in Australia.** In *East Malvern*. 4th edition. Victoria: Australian Breastfeeding Association; 2012:13.
59. Morton J: **The importance of hands.** *J Hum Lact* 2012, **28**:276–277.
60. Meek JY: **Breastfeeding in the workplace.** *Pediatr Clin North Am* 2001, **48**(2):461–474.

doi:10.1186/1471-2393-13-212

Cite this article as: Johns *et al.*: Prevalence and outcomes of breast milk expressing in women with healthy term infants: a systematic review. *BMC Pregnancy and Childbirth* 2013 **13**:212.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



Primary Care Interventions to Promote Breastfeeding: U.S. Preventive Services Task Force Recommendation Statement

U.S. Preventive Services Task Force*

Description: Update of a 2003 U.S. Preventive Services Task Force (USPSTF) recommendation on counseling to promote breastfeeding.

Methods: The USPSTF evaluated the results of a systematic review, conducted by the Tufts-New England Medical Center Evidence-based Practice Center, of literature published since January 2007 on primary care–initiated, –conducted, or –referable activities to promote and support breastfeeding.

Recommendation: The USPSTF recommends interventions during pregnancy and after birth to promote and support breastfeeding (Grade B recommendation).

Ann Intern Med. 2008;149:560-564.

www.annals.org

For author affiliation, see end of text.

* For a list of Task Force members, see the **Appendix**, available at www.annals.org.

The U.S. Preventive Services Task Force (USPSTF) makes recommendations about preventive care services for patients without recognized signs or symptoms of the target condition.

It bases its recommendations on a systematic review of the evidence of the benefits and harms and an assessment of the net benefit of the service.

The USPSTF recognizes that clinical or policy decisions involve more considerations than this body of evidence alone. Clinicians and policymakers should understand the evidence but individualize decision making to the specific patient or situation.

SUMMARY OF RECOMMENDATION AND EVIDENCE

The USPSTF recommends interventions during pregnancy and after birth to promote and support breastfeeding. This is a grade B recommendation.

See the **Figure** for a summary of the recommendation and suggestions for clinical practice.

Table 1 describes the USPSTF grades, and **Table 2** describes the USPSTF classification of levels of certainty

about net benefit. Both are also available at www.annals.org.

RATIONALE

Importance

There is convincing evidence that breastfeeding provides substantial health benefits for children and adequate evidence that breastfeeding provides moderate health benefits for women.

Effectiveness of Interventions to Change Behavior

Adequate evidence indicates that interventions to promote and support breastfeeding increase the rates of initiation, duration, and exclusivity of breastfeeding.

Harms of Interventions

No published studies focus on the potential direct harms from interventions to promote and support breastfeeding. The review did not include a search for potential harms of breastfeeding itself. The USPSTF has bounded the potential harms of interventions to promote and support breastfeeding as no greater than small.

USPSTF Assessment

The USPSTF concludes that there is moderate certainty that interventions to promote and support breastfeeding have a moderate net benefit.

CLINICAL CONSIDERATIONS

Patient Population under Consideration

This recommendation applies to pregnant women, new mothers, and young children. In rare circumstances involving health issues in mothers or infants, such as HIV infection or galactosemia, breastfeeding may be contraindicated and interventions to promote breastfeeding may not be appropriate. Interventions to promote and support

See also:

Print

Related article. 565

Summary for Patients. I-52

Web-Only

Appendix

CME quiz

Conversion of graphics into slides

Downloadable recommendation summary



breastfeeding may also involve a woman's partner, other family members, and friends.

Interventions

The current literature does not allow assessment of the individual aspects of multicomponent interventions or comparative effectiveness assessments of single-component interventions. The promotion and support of breastfeeding may be accomplished through interventions over the course of pregnancy; around the time of delivery; and after birth, while breastfeeding is under way. Interventions may include multiple strategies, such as formal breastfeeding education for mothers and families, direct support of mothers during breastfeeding observations, and the training of health professional staff about breastfeeding and techniques for breastfeeding support. Evidence suggests that interventions that include both prenatal and postnatal components may be the most effective at increasing breastfeeding duration. Many successful programs include peer support, prenatal breastfeeding education, or both.

OTHER CONSIDERATIONS

Implementation

Although the activities of individual clinicians to promote and support breastfeeding are likely to be positive, additional benefit may result from efforts that are integrated into systems of care. System-level interventions can incorporate clinician and team member training and policy development, and through senior leadership support and institutionalization, these initiatives may be more likely to be sustained over time. Although outside the scope of this recommendation and evidence review, community-based interventions to promote and support breastfeeding, such as direct peer-to-peer support, social marketing initiatives, workplace initiatives, and public policy actions, may offer additional sizeable benefits.

Research Needs and Gaps

Additional research is needed to better understand the effects of health care–based interventions to promote and support breastfeeding in the United States. Future research should include data collection on exclusive breastfeeding rates in addition to partial breastfeeding rates. Studies will be more useful if they are designed to allow some assessment of the relative contributions of individual components of multicomponent breastfeeding support programs. Research on the costs and cost–benefits of interventions is also needed. Additional research is needed to allow the tailoring of interventions to the needs of individual women and families. Good-quality prospective studies are needed to understand the effectiveness of compliance with the World Health Organization's Baby-Friendly Hospital Initiative in the United States, the contributions of individual components, and the interactive effect of the components with particular focus on postdischarge breastfeeding support.

DISCUSSION

Health Effects

In 2005, 73% of new mothers initiated breastfeeding, nearly reaching the U.S. Healthy People 2010 goal of 75% (1, 2). Thirty-nine percent breastfed their children for at least 6 months and 20% did so for 12 months (1). Fourteen percent of infants were exclusively breastfed for their first 6 months, as recommended by the American Academy of Pediatrics (AAP), the American Academy of Family Physicians (AAFP), and the U.S. Surgeon General (3–5).

Not breastfeeding is associated with health risks for mothers and children. For infants, not being breastfed is associated with increased numbers of ear infections, lower respiratory tract infections, and gastrointestinal infections (6). Children who were not breastfed were more likely to have asthma, type 2 diabetes, and obesity (6). For women, not breastfeeding is associated with higher rates of both breast and ovarian cancer (6).

Scope of Review

This recommendation is supported by a systematic evidence review conducted for the USPSTF by the Tufts-New England Medical Center Evidence-based Practice Center (7). The review updates the USPSTF's 2003 evidence report (8) and includes literature published between January 2001 and January 2007. Although the investigators included multiple study designs in their search strategies, the final report focused on randomized, controlled trials. The investigators limited studies to those with a focus on healthy term and near-term infants, their mothers, and members of the mother–child support team. As directed by the USPSTF, they used a broad conception of primary care interventions that encompassed activities initiated, conducted, or referable by primary care clinicians. Settings included primary care offices; labor, delivery, and postpartum inpatient settings; and patient homes. The review did not address community-based interventions, such as media campaigns, worksite lactation programs, and peer-to-peer support programs that do not interact with the health system.

Effectiveness of Interventions to Change Behavior

In evaluating more than 25 randomized trials of interventions conducted in the United States and in developed countries around the world, the USPSTF concluded that adequate evidence indicates that coordinated interventions throughout pregnancy, birth, and infancy can increase breastfeeding initiation, duration, and exclusivity. A large cluster randomized study of an intervention conducted in Belarus and modeled on the Baby-Friendly Hospital Initiative found that infants in the intervention group were significantly more likely than those in the control group to be exclusively breastfed and to have lower rates of gastrointestinal infections and atopic dermatitis (9). This good-quality study provides evidence of the potential effects of multifaceted breastfeeding interventions to improve health outcomes.

Potential Harms of Interventions

No studies identified for the USPSTF reported harms from interventions to promote and support breastfeeding. Nonetheless, there are potential harms, such as making women feel guilty. Breastfeeding interventions, like all other health care interventions designed to encourage healthy behaviors, should aim to empower individuals to make informed choices supported by the best available evidence. As with interventions to achieve a healthy weight or to quit smoking, breastfeeding interventions should be designed and implemented in ways that do not make women feel guilty when they make an informed choice not to breastfeed.

Estimate of Magnitude of Net Benefit

The USPSTF found that the benefits of breastfeeding are substantial and that the benefits of multimodal interventions to promote and support breastfeeding are moderate. Although the evidence was inadequate to determine the potential harms of these interventions, the USPSTF estimated these potential harms to be no greater than small. The USPSTF concluded with moderate certainty that the net benefits are moderate for multifaceted interventions to promote and support breastfeeding.

RECOMMENDATIONS OF OTHERS

The AAP, AAFP, and the American College of Obstetricians and Gynecologists all recommend that pregnant women receive breastfeeding education and counseling (3, 10, 11). The AAFP and AAP also recommend that peripartum policies and practices support breastfeeding mothers and infants and that breastfeeding families receive ongoing breastfeeding support (3, 10).

From the U.S. Preventive Services Task Force, Agency for Healthcare Research and Quality, Rockville, Maryland.

Disclaimer: Recommendations made by the USPSTF are independent of the U.S. government. They should not be construed as an official position of the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.

Financial Support: The USPSTF is an independent, voluntary body. The U.S. Congress mandates that the Agency for Healthcare Research and Quality support the operations of the USPSTF.

Potential Financial Conflicts of Interest: None disclosed.

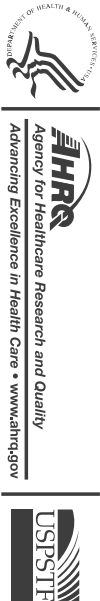
Requests for Single Reprints: Reprints are available from the USPSTF Web site (www.preventiveservices.ahrq.gov).

References

1. Breastfeeding Practices—Results from the National Immunization Survey, 2007. Atlanta: Centers for Disease Control and Prevention; 2007. Accessed at www.cdc.gov/breastfeeding/data/NIS_data/index.htm on 3 September 2008.
2. U.S. Department of Health and Human Services. Healthy People 2010: Conference Edition. Washington, DC: U.S. Government Printing Office; 2000.
3. Gartner LM, Morton J, Lawrence RA, Naylor AJ, O'Hare D, Schanler RJ, et al. American Academy of Pediatrics Section on Breastfeeding. Breastfeeding and the use of human milk. *Pediatrics*. 2005;115:496-506. [PMID: 15687461]
4. American Academy of Family Physicians. Breastfeeding (Policy Statement). Leawood, KS: American Acad Family Physicians; 2007. Accessed at www.aafp.org/online/en/home/policy/policies/b/breastfeedingpolicy.html on 9 January 2008.
5. Breastfeeding—Best For Baby. Best For Mom. Washington, DC: Department of Health and Human Services; 2007. Accessed at www.4women.gov/breastfeeding on 3 September 2008.
6. Ip S, Chung M, Raman G, Chew P, Magula N, DeVine D, et al. Breastfeeding and Maternal and Infant Health Outcomes in Developed Countries. Evidence Report/Technology Assessment no. 153. (Prepared by Tufts-New England Medical Center Evidence-based Practice Center, under contract no. 290-02-0022.) AHRQ Publication no. 07-E007. Rockville, MD: Agency for Healthcare Research and Quality; 2007. Accessed at www.ahrq.gov/downloads/pub/evidence/pdf/brfout/brfout.pdf on 3 September 2008.
7. Chung M, Raman G, Trikalinos T, Lau J, Ip S. Interventions in primary care to promote breastfeeding: an evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2008;149:565-82.
8. Guise JM, Palda V, Westhoff C, Chan BK, Helfand M, Lieu TA. U.S. Preventive Services Task Force. The effectiveness of primary care-based interventions to promote breastfeeding: systematic evidence review and meta-analysis for the US Preventive Services Task Force. *Ann Fam Med*. 2003;1:70-8. [PMID: 15040435]
9. Kramer MS, Chalmers B, Hodnett ED, Sevkovskaya Z, Dzikovich I, Shapiro S, et al. PROBIT Study Group (Promotion of Breastfeeding Intervention Trial). Promotion of Breastfeeding Intervention Trial (PROBIT): a randomized trial in the Republic of Belarus. *JAMA*. 2001;285:413-20. [PMID: 11242425]
10. American Academy of Family Physicians. Breastfeeding (Position Paper). Leawood, KS: American Acad Family Physicians; 2007. Accessed at www.aafp.org/online/en/home/policy/policies/b/breastfeedingpositionpaper.html on 9 January 2008.
11. American Academy of Pediatrics and American College of Obstetricians and Gynecologists. Guidelines for Perinatal Care. 6th ed. Elk Grove, IL: American Acad Pediatrics; 2007.

Figure. Primary care interventions to promote breastfeeding: clinical summary of a U.S. Preventive Services Task Force recommendation.

Annals of Internal Medicine



PRIMARY CARE INTERVENTIONS TO PROMOTE BREASTFEEDING:
CLINICAL SUMMARY OF U.S. PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

Population	Pregnant Women	New Mothers	The Mother's Partner, Other Family Members, and Friends	Infants and Young Children
Recommendation	Promote and support breastfeeding Grade: B			

Benefits of Breastfeeding	Mothers: Less likelihood of breast and ovarian cancer	Infants: Fewer ear, lower respiratory tract, or gastrointestinal infections	Young children: Less likelihood of asthma, type 2 diabetes, or obesity
Interventions to Promote Breastfeeding	Interventions to promote and support breastfeeding have been found to increase the rates of initiation, duration, and exclusivity of breastfeeding. Consider multiple strategies, including: • Formal breastfeeding education for mothers and families • Direct support of mothers during breastfeeding • Training of primary care staff about breastfeeding and techniques for breastfeeding support • Peer support Interventions that include both prenatal and postnatal components may be most effective at increasing breastfeeding duration. In rare circumstances, for example, for mothers with HIV or infants with galactosemia, breastfeeding is not recommended. Interventions to promote breastfeeding should empower individuals to make informed choices supported by the best available evidence.		
Implementation	System-level interventions with senior leadership support may be more likely to be sustained over time.		

For a summary of the evidence systematically reviewed in making these recommendations, the full recommendation statement, and supporting documents, please go to www.preventiveservices.ahrq.gov.

Table 1. What the USPSTF Grades Mean and Suggestions for Practice

Grade	Definition	Suggestions for Practice
A	The USPSTF recommends the service. There is high certainty that the net benefit is substantial.	Offer/provide this service.
B	The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial.	Offer/provide this service.
C	The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is moderate or high certainty that the net benefit is small.	Offer/provide this service only if other considerations support offering or providing the service in an individual patient.
D	The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits.	Discourage the use of this service.
I statement	The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.	Read clinical considerations section of USPSTF Recommendation Statement. If the service is offered, patients should understand the uncertainty about the balance of benefits and harms.

USPSTF = U.S. Preventive Services Task Force.

Table 2. U.S. Preventive Services Task Force Levels of Certainty Regarding Net Benefit

Level of Certainty*	Description
High	The available evidence usually includes consistent results from well-designed, well-conducted studies in representative primary care populations. These studies assess the effects of the preventive service on health outcomes. This conclusion is therefore unlikely to be strongly affected by the results of future studies.
Moderate	The available evidence is sufficient to determine the effects of the preventive service on health outcomes, but confidence in the estimate is constrained by such factors as: the number, size, or quality of individual studies inconsistency of findings across individual studies limited generalizability of findings to routine primary care practice lack of coherence in the chain of evidence. As more information becomes available, the magnitude or direction of the observed effect could change, and this change may be large enough to alter the conclusion.
Low	The available evidence is insufficient to assess effects on health outcomes. Evidence is insufficient because of: the limited number or size of studies important flaws in study design or methods inconsistency of findings across individual studies gaps in the chain of evidence findings that are not generalizable to routine primary care practice a lack of information on important health outcomes. More information may allow an estimation of effects on health outcomes.

*The U.S. Preventive Services Task Force (USPSTF) defines *certainty* as “likelihood that the USPSTF assessment of the net benefit of a preventive service is correct.” The net benefit is defined as benefit minus harm of the preventive service as implemented in a general primary care population. The USPSTF assigns a certainty level based on the nature of the overall evidence available to assess the net benefit of a preventive service.

APPENDIX: MEMBERS OF THE U.S. PREVENTIVE SERVICES TASK FORCE

Members of the U.S. Preventive Services Task Force[†] are Ned Calonge, MD, MPH, *Chair* (Colorado Department of Public Health and Environment, Denver, Colorado); Diana B. Petitti, MD, MPH, *Vice Chair* (Keck School of Medicine, University of Southern California, Sierra Madre, California); Thomas G. DeWitt, MD (Children's Hospital Medical Center, Cincinnati, Ohio); Allen Dietrich, MD (Dartmouth Medical School, Lebanon, New Hampshire); Kimberly D. Gregory, MD, MPH (Cedars-Sinai Medical Center, Los Angeles, California); Russell Harris, MD, MPH (University of North Carolina School of Medicine, Chapel Hill, North Carolina); George Isham, MD, MS (HealthPartners, Minneapolis, Minnesota); Michael L. LeFevre, MD, MSPH (University of Mis-

souri School of Medicine, Columbia, Missouri); Rosanne Leipzig, MD, PhD (Mount Sinai School of Medicine, New York, New York); Carol Loveland-Cherry, PhD, RN (University of Michigan School of Nursing, Ann Arbor, Michigan); Lucy N. Marion, PhD, RN (School of Nursing, Medical College of Georgia, Augusta, Georgia); Virginia A. Moyer, MD, MPH (University of Texas Health Science Center, Houston, Texas); Judith K. Ockene, PhD (University of Massachusetts Medical School, Worcester, Massachusetts); George F. Sawaya, MD (University of California, San Francisco, San Francisco, California); and Barbara P. Yawn, MD, MSPH, MSc (Olmsted Medical Center, Rochester, Minnesota).

[†]This list includes members of the Task Force at the time this recommendation was finalized. For a list of current Task Force members, go to www.ahrq.gov/clinic/uspstfab.htm.



Center for Medicaid and CHIP Services

Medicaid Coverage of Lactation Services

Issue

This issue brief sets forth current levels of State Medicaid coverage for lactation services and explores how CMS can encourage and assist States in increasing access to such services.

Background

Improving the health of the population and reducing preventable causes of poor health, such as obesity, is a priority of the Department of Health and Human Services (HHS) and the Centers for Medicare & Medicaid Services (CMS).ⁱ Current research shows that the practice of breastfeeding for the first 6 to 12 months of life is highly beneficial for both the mother and infant. On January 20, 2011, the United States Surgeon General released “The Surgeon General’s Call to Action to Support Breastfeeding.”ⁱⁱ This report indicates that there is a 32% higher risk of childhood obesity and a 64% higher risk of type 2 diabetes for children who are not breastfed. An extensive body of research supports these assertions and provides evidence of the positive effects of breastfeeding on both short and long term infant and maternal health.ⁱⁱⁱ Professional organizations advocate for exclusive breastfeeding during the first 6 months of life, meaning that infants should not be given any other substance other than breast milk, including water.^{iv} Breastfeeding also serves additional advantages for low birth weight infants. Human milk consumption is associated with a reduction in sepsis infections and gastrointestinal illnesses among very low birth weight newborns in the neonatal intensive care unit (NICU).^v

The U.S. Preventive Services Task Force (USPSTF) specifically recommends coordinated interventions throughout pregnancy, birth, and infancy to increase breastfeeding initiation, duration, and exclusivity.^{vi} Such recommended interventions include formal breastfeeding education for mothers and families, direct support of the mother during breastfeeding observations, and the training of health professional staff about breastfeeding and techniques for breastfeeding support.^{vii} The opportunity to promote and provide support for breastfeeding occurs many times throughout the interactions that women have with various types of health care providers during and after pregnancy.

Categories of Coverage for Lactation Services

All States participating in the Medicaid program cover pregnancy-related services. 42 C.F.R. § 440.210 defines “pregnancy-related services” as those that are necessary for the health of the pregnant woman and fetus, or that have become necessary as a result of the woman having been pregnant. These include, but are not limited to, prenatal care, delivery, postpartum care, and family planning services. States must provide coverage of pregnancy-related services for an

extended postpartum period, defined as beginning on the last day of pregnancy and extending through the end of the month in which the 60-day period following termination of pregnancy ends. This definition of “pregnancy-related services” is broad enough to encompass lactation services.

Due to the multiple health benefits associated with breastfeeding, CMS encourages States to go beyond the requirement of solely coordinating and referring enrollees to the Special Supplemental Food Program for Women, Infants, and Children (WIC) (established in 42 C.F.R. § 431.635) and include lactation services as separately reimbursed pregnancy-related services.

States may use the following Medicaid coverage categories to reimburse lactation services:

- Inpatient hospital services (other than services in an institution for mental disease), per Social Security Act (SSA) § 1905(a)(1);
- Outpatient hospital services, per SSA § 1905(a)(2)(A) and 42 C.F.R. § 440.10;
- Early and periodic screening, diagnostic, and treatment services for individuals who are eligible under the plan and are under the age of 21, per SSA § 1905(a)(4)(B);
- Physicians’ services furnished by a physician under the physician’s supervision, whether furnished in the office, the patient’s home, a hospital, or a nursing facility, or elsewhere, per SSA § 1905(a)(5)(A);
- Services furnished by a nurse-midwife, which the nurse-midwife is legally authorized to perform under State law, per SSA § 1905(a)(17);
- Freestanding birth center services, per SSA § 1905(a)(28); and
- Services furnished by nurse practitioners per 42 C.F.R. § 440.166 and other licensed practitioners per 42 C.F.R. § 440.60.

Current State Practices

Because lactation services are not specifically mentioned in the Medicaid statute or Federal Medicaid regulations, not all States separately reimburse lactation services as pregnancy-related services.^{viii} In fact, States vary widely in the amount and scope of coverage they provide for lactation services as a part of prenatal, postpartum, and infant care.

The Kaiser Family Foundation and the George Washington University Department of Health Policy conducted a survey in 2008 to assess each State’s Medicaid coverage of perinatal services.^{ix} This survey, to which 44 States provided a response, focused on three main categories of lactation services: 1) Breastfeeding Education, 2) Individual Lactation Consultation, and 3) Equipment Rentals (see attachment). The findings of the study showed that despite the establishment of breastfeeding as an important preventive health measure, Medicaid coverage of lactation services was far from comprehensive. While States do cover lactation related services as a part of the hospital/facility fee after delivery, not all States cover and separately reimburse for such services. Specifically, 25 of the responding States covered breastfeeding education services, 15 of the responding States covered individual lactation consultations, and 31 of the responding States covered equipment rentals.

Some States have taken additional or alternative steps to promote breastfeeding. For example, some provide Medicaid coverage for donor milk from a human milk bank for infants whose mothers are unable to breastfeed due to medical reasons or in cases where the infant cannot tolerate formula use.^x New Hampshire Medicaid provides coverage of a Maternal Postpartum Assessment, which specifically includes an evaluation of whether the mother is properly breastfeeding and provides the accompanying postpartum education to ensure she continues to breastfeed.^{xi} Florida operates a Medicaid waiver program that coordinates prenatal care through monthly outreach and case management. The program specifically stresses healthy nutrition and breastfeeding habits, but is limited to Medicaid enrollees identified as high-risk for poor birth outcomes or those referred to the program by their health care provider.^{xii}

Aside from variation in the amount and scope of coverage provided for lactation services, States also differ in the ways that these services are billed and coded. For example, in Illinois, both breastfeeding education and individual lactation consultations are billed as part of an exam, not as a separate service.^{xiii} Some States simply allow providers to bill individual lactation consultations as part of an “Evaluation and Management” visit. In many cases, lactation consultations may be provided as part of Childbirth Education Classes or covered (and coded) as part of a woman’s inpatient hospital stay.^{xiv} The home health nurse visit provided after discharge from the hospital also allows health care providers to engage in a one-on-one lactation evaluation and consultation for new mothers.

The following provides examples of different codes that States use for billing and receiving federal matching funds for coverage of lactation services:

- Lactation Consultation (face-to-face visit), HCPCS code S9443
- Postpartum Care and Examination of Lactating Mother, ICD-9 code V24.1
- Manual Breast Pump purchase, CPT Code E0602
- Hospital Grade Electric Breast Pump rental, CPT Code E0604
- Individual Electric Breast Pump purchase, CPT Code E0603

Example of a State Benefit Package

Rhode Island provides the following benefit package for breastfeeding mothers enrolled in Medicaid. None of the services are associated with co-payments.^{xv}

	Benefit	Criteria
Education	Prenatal Breastfeeding Classes/ Childbirth Education Classes	Covered benefit – group and individual sessions No referral or authorization needed
	Breastfeeding Support Group	Not a covered benefit
Lactation Support	In-patient hospital	Covered as part of inpatient stay
	Outpatient hospital	Covered benefit – initial consult must occur within first two weeks of delivery. Benefit limited to 2 additional visits within first month after delivery

	Home	Covered benefit – prior approval required. Initial consult must occur within first two weeks of delivery. Benefit limited to 2 additional visits within first month after delivery
Equipment	Manual Breast Pump Purchase	Covered benefit for medical necessity or for mother returning to work or school (up to child's first birthday) Requires prescription
	Hospital Grade Electric Breast Pump Rental	Covered benefit for medical necessity or for mother returning to work or school (up to child's first birthday) Requires prescription
	Individual Electric Breast Pump Purchase	Covered benefit for medical necessity or for mother returning to work or school (up to child's first birthday) Requires prescription
	Pump Kits Purchase	Covered benefit – authorized with electric pump (1 maximum)

Options for Improving Access to Lactation Services

Aside from providing a benefit package, there are a number of strategies that States can undertake to improve access to lactation services for Medicaid beneficiaries.

- Encourage managed care entities/organizations and other providers to direct patients to Baby-Friendly Hospitals. The Baby-Friendly Hospital Initiative was established by the World Health Organization (WHO) and UNICEF in 1991 and gives special recognition to hospitals and birthing centers that follow the Ten Steps to Successful Breastfeeding for Hospitals.^{xvi} As of November 17, 2011, there are 121 Baby-Friendly Hospitals and Birth Centers throughout the United States that provide new mothers with the information and skills needed to successfully initiate and continue breastfeeding.^{xvii} Baby-Friendly Hospitals have experienced significant increases in the rates of successful and sustained breastfeeding among Medicaid and privately insured mothers, more so than in non-Baby-Friendly Hospitals.^{xviii}
- Because primary care interventions to promote and support breastfeeding received a grade of B from the USPSTF, coverage of lactation services without cost-sharing will be eligible for a 1 percentage point increase in federal medical assistance percentage (FMAP) in 2013 per Section 4106 of the Affordable Care Act. CMS encourages States to begin formulating coverage policies for such services.

- Encourage managed care organizations to either collect data that establishes a baseline for breastfeeding rates and/or initiate a performance improvement project (PIP) or focused study that seeks to increase rates of breastfeeding within the population. For example, a health plan in New York State operates a performance improvement project that seeks to increase education among pregnant women, mothers, and providers while specifically stressing breastfeeding in the first years of life and provider encouragement of breastfeeding.^{xix}
- Eliminate variation of coverage among Medicaid managed care plans and encourage managed care contractors to provide breastfeeding education, either by means of referrals to Special Supplemental Food Program for Women, Infants, and Children (WIC) or direct provision of the services through its provider network.^{xx} WIC provides breastfeeding women with lactation support by means of counseling, the provision of breastfeeding educational materials, the option to receive an enhanced food package, follow-up support through peer counselors, and the provision of breast pumps or other equipment to help reinforce a healthy and successful breastfeeding routine.^{xxi} ^{xxii}
- Coordinate with the “*Let’s Move!*” initiative, launched by the First Lady of the United States, Michelle Obama, in an effort to promote breastfeeding as one of the many strategies to prevent childhood obesity.^{xxiii}
- Disseminate information to providers that discuss the benefits of lactation services, as well as information and other resources regarding the promotion of lactation services.

January 10, 2012

ⁱ Other agencies within HHS have undertaken the following initiatives related to breastfeeding: 1) The Centers for Disease Control and Prevention (CDC) publishes a CDC Guide to Breastfeeding Interventions (Shealy KR, Li R, Benton-Davis S, Grummer-Strawn LM. The CDC Guide to Breastfeeding Interventions. Atlanta: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, 2005) ; 2) The Health Resources and Services Administration (HRSA) operates the Business Case for Breastfeeding program, which is a comprehensive program designed to educate employers about the value of supporting breastfeeding employees in the workplace; 3) The Agency for Healthcare Research and Quality (AHRQ) provides information and research that assesses breastfeeding with maternal and infant health outcomes; 4) The Office on Women’s Health published the 2000 HHS Blueprint for Action on Breastfeeding.

ⁱⁱ US Department of Health and Human Services, The Surgeon General’s Call to Action to Support Breastfeeding, 2011, available at: <http://www.surgeongeneral.gov/topics/breastfeeding/calltoactiontosupportbreastfeeding.pdf>.

ⁱⁱⁱ Ip S, et al. Breastfeeding and Maternal and Infant Outcomes in Developed Countries: Evidence Report/Technology Assessment Number 153, Agency for Health Care Research and Quality Publication No. 07-E007; April 2007, available at: <http://www.ahrq.gov/downloads/pub/evidence/pdf/brfout/brfout.pdf>.

^{iv} The American Academy of Pediatrics, American Academy of Family Physicians, American College of Obstetricians and Gynecologists, American College of Nurse Midwives, American Dietetic Association, and American Public Health Association all recommend that infants be breastfed for a minimum of 12 months.

^v El-Mohandes A, Picard MB, Simmens SJ, et al. Use of human milk in the intensive care nursery decreases the incidence of nosocomial sepsis *Journal of Perinatology* 1997; 2:130-134; Taylor G, Minich N, Hack, M. The Effect of Maternal Milk on Neonatal Morbidity of Very Low-Birth-Weight Infants. *Arch Pediatric and Adolescent Medicine*. 2003;157:66-71; Meinen-Derr J, Poindexter B, Wrage, L, Morrow A, Stoll B, Donovan E. Role of human milk in extremely low birth weight infants’ risk of necrotizing enterocolitis or death, *Journal of Perinatology*. 2009 Jan; 29(1):57-62.

^{vi} USPSTF, Primary Care Interventions to Promote Breastfeeding: Recommendation Statement, October 2008, available at: <http://www.uspreventiveservicestaskforce.org/uspstf08/breastfeeding/brfeeds.htm>.

^{vii} Ibid.

^{viii} See 42 C.F.R. 440.210 Required Services for the Categorically Needy and 42 C.F.R. 440.220 Required Services for the Medically Needy.

^{ix} Stewart A, Cox M, Doamekpor L, Ranji U, Salgancioff A, *State Medicaid Coverage of Perinatal Services: Summary of State Survey Findings*. The Henry J. Kaiser Family Foundation, November 2009, available at: <http://www.kff.org/womenshealth/upload/8014.pdf>.

^x Texas Medicaid and Health Partnership, Texas Health Steps Medicaid Manual: Donor Human Milk, available at: <http://www.tmhp.com/HTMLmanuals/TMPMP/2008/2008%20TMPMP-46-108.html>.

^{xi} N.H. ADMIN. RULES [HE-W] 547.04 (2011).

^{xii} AHCA, Florida Medicaid Summary of Services for Fiscal Year 2010-2011 at 105, available at http://ahca.myflorida.com/Medicaid/pdf/SS_10_100501_SOS_ver2-4_1164_1011_FINAL2.pdf.

^{xiii} Stewart A, Cox M, Doamekpor L, Ranji U, Salgancioff A, *State Medicaid Coverage of Perinatal Services: Summary of State Survey Findings*. The Henry J. Kaiser Family Foundation, November 2009, available at: <http://www.kff.org/womenshealth/upload/8014.pdf>.

^{xiv} Rhode Island Department of Health, *Breastfeeding Insurance Benefits Guidelines*, available at: www.health.ri.gov/family/breastfeeding/insurancebenefits.php.

^{xv} Ibid.

^{xvi} The Ten Steps to Successful Breastfeeding are: 1) Have a written breastfeeding policy that is routinely communicated to all health care staff; 2) Train all health care staff in skills necessary to implement this policy; 3) Inform all pregnant women about the benefits and management of breastfeeding; 4) Help mothers initiate breastfeeding within one hour of birth; 5) Show mothers how to breastfeed and how to maintain lactation, even if they are separated from their infants; 6) Give newborn infants no food or drink other than breast milk, unless *medically* indicated; 7) Practice “rooming in” which is allowing the mother and infant to remain together 24 hours a day; 8) Encourage breastfeeding on demand; 9) Give no pacifiers or artificial nipples to breastfeeding infants; and 10) Foster the establishment of breastfeeding support groups and refer mothers to them on discharge from the hospital or clinic. For more information, please see <http://www.babyfriendlyusa.org/eng/index.html>.

^{xvii} To be recognized as a Baby-Friendly Hospital, facilities must register with Baby-Friendly USA and undergo an on-site assessment to demonstrate the facility has integrated the “Ten Steps to Successful Breastfeeding” into hospital policies and procedures for healthy newborns. More detailed information can be found in the Guidelines and Evaluation Criteria for Facilities Seeking Baby-Friendly Designation, 2010 edition, available at: http://babyfriendlyusa.org/eng/docs/2010_Guidelines_Criteria_4.19.11.pdf.

^{xviii} Philipp B, Merewood A, Miller L, et al., Baby-Friendly Hospital Initiative Improves Breastfeeding Initiation Rates in a US Hospital Setting. *Pediatrics*. 2001; 108(3): 677-681.

^{xix} New York State Medicaid Managed Care Performance Improvement Projects, 2009-2010 Pediatric Obesity - Summary of Projects, available at: http://www.health.ny.gov/health_care/managed_care/reports/pediatric_obesity.htm.

^{xx} Specifically, 42 C.F.R. 431.635 requires the coordination of Medicaid with WIC, which includes providing written notice of the availability of WIC benefits to all individuals in the State who are determined eligible for Medicaid and who are pregnant, postpartum, breastfeeding, or a child under the age of five. In addition, the State must at least annually provide written notice of the availability of WIC benefits (including the telephone number and address of the local WIC agency) to all Medicaid recipients who might be pregnant, postpartum, or breastfeeding.

^{xxi} United States Department of Agriculture, Food & Nutrition Service, Women, Infants, and Children, available at: <http://www.fns.usda.gov/wic/>.

^{xxii} All states that utilize managed care to deliver Medicaid services are required by 42 C.F.R. 438.358 to engage in three mandatory external quality review activities, including the validation of PIPs and performance measures.

^{xxiii} For more information, please see www.letsmove.gov.

HERC Guideline Information, with References

Problem statement:

Despite Oregon's high initiation rates, breastfeeding drops off in the early weeks and months, even when mothers' initial plans are to breastfeed for longer. Many of these mothers report either not having access to lactation services or not being able to access services at the time it was needed. Many women on OHP are routinely denied lactation services once discharged from the hospital.

Why provide lactation services? What is the evidence base for this recommendation?

The United States Preventive Services Task Force (USPSTF) recommends coordinated interventions during pregnancy and after birth to promote and support breastfeeding in order to increase the rates of initiation, duration and exclusivity. It is given a Grade B.¹

The American Academy of Pediatrics (AAP)², the American Academy of Family Physicians (AAFP)³, and the American Congress of Obstetricians and Gynecologists (ACOG)⁴ support exclusive breastfeeding for the first six months of life, followed by breastfeeding in combination with the introduction of complementary foods until at least 12 months of age and thereafter for as long as mutually desired by mother and baby. All three organizations recommend that pregnant women receive breastfeeding education and counseling throughout the perinatal period.

The Health Resources and Services Administration (HRSA)⁵ supports the Institute of Medicine's⁶ recommendations on preventive services that address health needs specific to women and fill gaps in existing guidelines.

A 2012 review found evidence that formal breastfeeding education, as well as peer counseling and lactation consultation, during pregnancy appear to increase breastfeeding duration⁷. Furthermore, evidence suggests that interventions with combined prenatal and postnatal components might be the most effective way to increase breastfeeding duration.^{8,9}

Women are significantly more likely to achieve their breastfeeding goals if they are supported prenatally, in the maternity care facility, and then after discharge. A small investment in lactation care and services early in a child's life reaps a long-term positive return on investment. Implementation of Medicaid coverage for lactation counseling services would be expected to result in an increase in breastfeeding initiation, exclusivity and duration consistent with the recommendations of expert groups and the USPSTF. Besides the short- and long-term health benefits to mothers and infants from breastfeeding, there will be health care cost savings as well.¹⁰

The national Centers for Medicare and Medicaid Services (CMS) also recommend the provision of lactation services. Federal guidance entitled "Medicaid Coverage of Lactation Services",

issued January 10, 2012, states “Due to the multiple health benefits associated with breastfeeding, CMS encourages States to go beyond the requirement of solely coordinating and referring enrollees to the Special Supplemental Food Program for Women, Infants, and Children (WIC) and **include lactation services as separately reimbursed pregnancy-related services.**”¹¹

In fact, effective January 1, 2013, states became eligible for an increased federal medical assistance percentage if they covered such preventive services, including breastfeeding counseling, with no co-pay (ACA § 4106). States have the flexibility to cover breastfeeding services in numerous different ways under Medicaid, including inpatient and outpatient hospital services.^{1, 10, 12}

Which lactation services are recommended?

- Comprehensive lactation support and counseling services are recommended^{5, 6}, including interventions during pregnancy, as well as after birth. Services to be provided include:
 - Prenatal breastfeeding education groups
 - Postnatal breastfeeding education and support groups
 - Individual consultations with a lactation consultant – both inpatient and outpatient – to address specific breastfeeding challenges¹⁴
- The provision of breastfeeding equipment, including high quality electric breast pumps, are also recommended as needed to ensure that breastfeeding is well established, and that a woman’s milk production is protected and maintained.¹⁴
- Access to lactation services and equipment would occur through referral from a woman’s health care provider.¹⁵
- Lactation services would be provided by a qualified provider – see section below.^{5, 14, 15}
- Items and services would be covered without cost-sharing, subject to reasonable medical management, which may include purchase instead of rental of equipment.^{5, 13}
- Coverage of comprehensive lactation support and counseling, and costs of renting or purchasing breastfeeding equipment, extends for the duration of breastfeeding.¹³

Who should provide reimbursable lactation services?

While qualified practitioners of Medicaid reimbursable lactation counseling services can include licensed, registered, or certified health care professionals such as physicians, nurse practitioners, midwives, physician assistants, registered nurses and licensed dietitians, these health care professionals often do not have the depth and breadth of training to provide higher level lactation services. **International Board Certified Lactation Consultants (IBCLCs)** credentialed by the International Board of Lactation Consultant Examiners (IBLCE) should be Medicaid reimbursable practitioners for lactation counseling services.

IBCLCs are specifically educated in human lactation and breastfeeding, have clinical practice in breastfeeding care, and are trained to deliver lactation services. They have been working in the health care field for 25 years in the U.S. and worldwide in 81 countries. An IBCLC is skilled to

encourage informed decision-making about infant feeding, develop/implement care plans, minimize the potential for breastfeeding problems or complications, facilitate referrals, and communicate with other health care professionals as appropriate.

For routine prenatal and postpartum breastfeeding education and support, particularly in group settings, peer providers such as trained traditional (community) health workers and doulas would be the recommended providers for this level of lactation services. In Oregon, these paraprofessionals are already Medicaid reimbursable providers.

The CMCS Informational Bulletin from November 27, 2013, outlines the change in federal regulations allowing reimbursement for non-licensed providers. This bulletin states that services must be “recommended by a physician or other licensed practitioner” when previously this regulation stated that “Preventive services means services provided by a physician or licensed practitioner...”¹⁵

References

1. US Preventive Services Task Force. Primary care interventions to promote breastfeeding: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med* 2008; 149: 560-4. Available at <http://www.uspreventiveservicestaskforce.org/uspstf/uspsbrfd.htm>
2. "Breastfeeding and the Use of Human Milk", Policy Statement, Section on Breastfeeding, American Academy of Pediatrics, published online May 4, 2012. Available at <http://pediatrics.aappublications.org/content/129/3/e827.abstract>
3. “Breastfeeding, Family Physicians Supporting” (Position Paper), 2014, American Academy of Family Physicians. Available at <http://www.aafp.org/about/policies/all/breastfeeding-support.html>
4. *ACOG Clinical Review* "Special Report From ACOG, Breastfeeding: Maternal and Infant Aspects", Number 361, February 2007 (Reaffirmed 2013), American Congress of Obstetricians and Gynecologists. Available at <http://www.acog.org/Resources-And-Publications/Committee-Opinions/Committee-on-Health-Care-for-Underserved-Women/Breastfeeding-Maternal-and-Infant-Aspects>
5. Health Resources and Services Administration. Women’s preventive services guidelines. Rockville, MD: US Department of Health and Human Services, Health Resources and Services Administration; 2014. Available at <http://www.hrsa.gov/womensguidelines/>
6. Institute of Medicine of the National Academies, Report Brief, July 2011. Available at http://www.iom.edu/~media/Files/Report%20Files/2011/Clinical-Preventive-Services-for-Women-Closing-the-Gaps/preventiveservicesforwomenreportbrief_updated2.pdf

7. Lumbiganon P, Martis R, Laopaiboon M, Festin MR, Ho JJ, Hakimi M. Antenatal breastfeeding education for increasing breastfeeding duration. Cochrane Database Syst Rev 2012; (9):CD006425.
8. Chung M, Raman G, Trikalinos T, Lau J, Ip S. Interventions in primary care to promote breastfeeding: an evidence review for the U.S. Preventive Services Task Force. Ann Intern Med 2008;149:565–82.
9. Britton C, McCormick FM, Renfrew MJ, Wade A, King SE. Support for breastfeeding mothers. Cochrane Database Syst Rev 2007; (1):CD001141.
10. Centers for Disease Control and Prevention, Morbidity and Mortality Weekly Report. Prenatal Breastfeeding Counseling – Pregnancy Risk Assessment Monitoring System, United States, 2010. Available at http://www.cdc.gov/mmwr/preview/mmwrhtml/su6302a3.htm?s_cid=su6302a3_w
11. Center for Medicaid and CHIP Services, “Medicaid Coverage of Lactation Services”, January 10, 2012. Available at http://www.medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Quality-of-Care/Downloads/Lactation_Services_IssueBrief_01102012.pdf
12. Letter from CMS to State Medicaid Director, Feb. 1, 2013. Available at <http://www.medicaid.gov/Federal-Policy-Guidance/downloads/SMD-13-002.pdf>
13. Centers for Medicare and Medicaid Services, Center for Consumer Information & Insurance Oversight, Affordable Care Act Implementation FAQs - Set 12, Questions 18 & 20. Available at http://www.cms.gov/CCIIO/Resources/Fact-Sheets-and-FAQs/aca_implementation_faqs12.html
14. “Model Policy – Payer Coverage of Breastfeeding Support and Counseling Services, Pumps and Supplies”, 2nd Revised Edition, 2014, United States Breastfeeding Committee & National Breastfeeding Center. Available at <http://www.usbreastfeeding.org/LegislationPolicy/ExistingLegislation/ModelPolicyPayerCoverage/tabid/344/Default.aspx>
15. Centers for Medicare and Medicaid Services, CMCS Informational Bulletin: “Update on Preventive Services Initiatives”, Nov. 27, 2013 – References revised 42CFR440.130(c) Available at <http://medicaid.gov/Federal-Policy-Guidance/Downloads/CIB-11-27-2013-Prevention.pdf>

MEDICAID COVERAGE—Lactation Services and Supplies

State Examples

CALIFORNIA

California's Medicaid program, Medi-Cal, has had regulations to include breastfeeding services and supplies since 1999. Coverage is defined for provider and location (e.g. RN, RD, under MD, or clinic). Now many third party Medi-Cal contractors are covering breastfeeding services at greater lengths with IBCLCs. Implementation by these Medi-Cal managed care plans and providers is still uneven as some counties are doing well providing timely services and others are doing poorly.

Most current managed care plan letter. Currently provider updates are in the Medi-Cal Provider Bulletins <http://www.dhcs.ca.gov/formsandpubs/Documents/MMCDAPLsandPolicyLetters/PL1998/MMCDPL98010.pdf>

Note: the TAR for pump rentals is actually not required until 60 days. It used to be shorter based on a calculation of rental costs.

Since 1986 donor human milk has been a benefit of the Medi-Cal program for outpatient clients. Hospitals cannot bill for donor milk. The San Jose milk bank also provides Medi-Cal breast pumps to outpatient families locally in Silicon Valley. These stripped down electrical pumps are not rentals, moms keep them. Acceptable diagnostic codes are fairly broad. Pump rentals are reserved to hospital grade pumps and can be provided for two weeks with authorization. After that time it has to be authorized regularly.

Comprehensive Perinatal Services Program (CPSP), a Medicaid program, can be utilized to cover breastfeeding services as well. In this program providers are reimbursed for health education, group as well as individual education, and breastfeeding is included.

http://www.cce.csus.edu/portal/admin/handouts/CPSP_ProviderHandbook_2014May15.pdf

Contact Karen Farley, California WIC Association, kfarley@calwic.org for more information about Medi-Cal coverage of breastfeeding.

DELAWARE

United Medicaid in Delaware covers 6 postnatal lactation visits and double electric breast pumps. Aetna has Delaware's largest Medicaid contract and does not cover lactation visits. They will only issue pumps every 3 years.

NEW YORK

In 2012, the New York State Department of Health (NYSDOH) created a state plan amendment requiring Medicaid to reimburse for evidence-based breastfeeding education and lactation counseling consistent with the United States Preventive Services Task Force (USPSTF) recommendation. To date, New York is one of the only states operationalizing the benefit included in the women's preventive service regulation in its Medicaid program.

NYS Medicaid provides reimbursement for lactation services:

https://www.health.ny.gov/community/pregnancy/breastfeeding/medicaid_coverage/lactation_counseling_services.htm

Note: They don't reimburse *all* IBCLCs, but do reimburse IBCLCs that are also health care professionals (MD, RN, NP, PA, etc.) They also have a clear agreement with WIC.

NYS Medicaid also provides reimbursement for breast pumps:

https://www.health.ny.gov/community/pregnancy/breastfeeding/medicaid_coverage/breast_pump_coverage.htm

Here is the link to their amendment and approval documents through CMS:

http://www.health.ny.gov/regulations/state_plans/status/non-inst/approved/docs/app_2012-12-28_spa_12-16.pdf

NORTH CAROLINA

In 2012, the North Carolina Perinatal Committee of the North Carolina Child Fatality Task Force (CFTF), which included the breastfeeding coordinator for the state, designed a benchmark definition of lactation support and counseling for IBCLC coverage under the ACA provision. That benchmark specifies that lactation services are to be provided by a “trained” clinician, defined as an individual with the IBCLC credential or a physician or a physician extender, which include either a nurse practitioner or physician’s assistant. It is important to note that this proposal is in the initial phases of development. (AMCHP)

OKLAHOMA

Oklahoma Medicaid has paid for outpatient lactation services since 2007. Benefit is on mom so ends when her coverage ends, roughly 60 days post-partum. Up to 6 visits – prenatal or postpartum – independently billed by an IBCLC who is a licensed nurse or dietitian. The hope is to switch coverage to the baby at birth so the coverage doesn’t end so soon. And also include non-licensed IBCLCs as providers.

<https://www.okhca.org/providers.aspx?id=11795>

<http://www.okhca.org/providers.aspx?id=11795&terms=lactation%20services>

NATIONAL EFFORTS & RESOURCES

AMCHP issue a brief showing exemplary ACA coverage examples:

<http://www.amchp.org/Policy-Advocacy/health-reform/Documents/AMCHP%20Dec%202013%20Breastfeeding%20Issue%20Brief%20FINAL.pdf>

They have recently been funded by WKKF to do this as a comprehensive review of the whole country.

National Women’s Law Center toolkit:

http://www.nwlc.org/sites/default/files/pdfs/final_nwlclogo_preventiveservicestoolkit_5-21-14.pdf

CMS is working on a document to delineate the Medicaid coverage by state, including pump coverage and services provided. It is anticipated that this resource will be available by the end of the year.

USLCA www.uslca.org

Per Judith L. Gutowski, BA, IBCLC, RLC, USLCA Licensure & Reimbursement Committee Chair
Containing Health Care Costs Help in Plain Sight, 3rd edition. Page 23 and 28 for New York and Oklahoma examples (2nd edition on web presently, not 3rd edition)

Cost-Benefit Rationale for Lactation Services in North Carolina (unable to find on web; pdf available)

The Affordable Care Act – <https://www.healthcare.gov/what-are-my-breastfeeding-benefits/>

Frequently Asked Questions – Oregon WIC’s Breast Pump Program

Q: Who qualifies to receive breast pumps from the Oregon WIC Program?

A: Breastfeeding women who are not able to obtain a breast pump through their health care provider/insurer may qualify for a pump if they are enrolled in the Oregon WIC Program and they meet the specific criteria required for the type of pump offered.

Q: What kinds of breast pumps are available through Oregon WIC?

A: There are 3 types of breast pumps that are offered by the Oregon WIC Program:

- *Multi-user double electric breast pumps* (Medela Lactinas) are loaned to breastfeeding mothers with medical needs.
- *Personal double electric breast pumps* (Medela Pump in Styles) are given to breastfeeding mothers who are returning to work or school.
- *Manual breast pumps* (Medela Harmony or 2-handed pumps) are provided for occasional or short-term pumping needs.

The type of pump is determined by a trained WIC staff member.

Q: What kinds of medical needs qualify a breastfeeding mother for a multi-user breast pump?

A: Multi-user breast pumps are loaned for medical circumstances such as:

- A premature infant
- An infant with severe feeding problems
- A mother or infant who is ill or hospitalized
- Multiple birth, such as twins or triplets
- A mother who has had breast surgery

This is not a comprehensive list. The need for a breast pump is determined by a hospital lactation consultant or a WIC staff member with expertise in breastfeeding.

Q: What criteria must be met for a mother to receive a personal double electric breast pump?

A: A breastfeeding mother must meet **all** of the following requirements:

- Mother is working or going to school more than 20 hours per week, is away from her baby for 9 or more feedings per week, and is unable to feed her baby while at work or at break time
- Infant is exclusively breastfeeding
- Infant is at least 4 weeks old
- Mother and baby are having no problems with breastfeeding
- Mother plans to breastfeed exclusively for at least 4-6 months and will not expect any formula from WIC

Q: In what circumstances are manual pumps issued to breastfeeding mothers?

- A:** Manual pumps may be given to mothers who are experiencing:
- Temporary problems such as engorgement or flat/inverted nipples
 - Brief separations from the infant for social events or meetings
 - Time away from the infant to seek employment
 - Working or going to school on a part-time basis (less than 20 hours per week)
 - Other short term separations, where pumping would not occur more than 8 times per week
 - No access to electricity when it is time to pump

Q: How does a mother obtain a breast pump from the Oregon WIC Program?

- A:** Breast pumps are available from all local WIC agencies in Oregon and may be issued to breastfeeding women who meet the qualifications. An appointment may be required due to the time it takes for assessment, issuance, pump education, and documentation.

Q: Is there a fee for a WIC participant to receive a breast pump?

- A:** No. Multi-user breast pumps are *loaned* to WIC participants who meet the criteria, so there is no fee involved. Personal double electric breast pumps and manual breast pumps are available at no cost to breastfeeding mothers who qualify.

Q: Are there circumstances when breast pumps may NOT be issued?

- A:** Yes. Breast pumps may not be issued to:
- Pregnant women
 - Breastfeeding women who do not meet the necessary qualifications
 - Breastfeeding women who are using medical marijuana or any illicit substance
 - Participants who have not returned a multi-user breast pump in the past

Q: Does Oregon WIC rent or loan breast pumps to non-WIC mothers?

- A:** No. Oregon WIC does not rent breast pumps, and non-WIC participants do not qualify for any type of breast pump (or loan) from the Oregon WIC Program.

Lactation Support Services

Cost/Benefit Rationale

- Breastfeeding rates for women on Medicaid in NC are well below the Healthy People 2020 goals. A primary reason for this disparity is that women get discouraged and stop breastfeeding when problems develop and they do not have access to care providers who offer lactation services.^{1,2} Medical lactation services are not currently covered by Medicaid in NC.
- The availability of lactation support services, particularly when provided by highly qualified International Board Certified Lactation Consultants (IBCLCs) is intended to bring North Carolina's rates up to the national goals. If this occurs, the benefits in terms of reduced infant/early childhood mortality and morbidity, as well as reduced medical care costs would be significant.

Breastfeeding Goal	Healthy People 2020 Target	2011 NC WIC Rates (proxy for Medicaid numbers) *
Initiation	81.9%	63%
6 Months	60.6%	23.2%
12 Months	34.1%	18.1%

- A systematic review and meta-analysis on breastfeeding and infant/early childhood outcomes concludes that breastfeeding reduces Sudden Infant Death Syndrome, gastrointestinal infection, upper and lower respiratory disease and hospitalization, childhood leukemia, asthma, otitis media, childhood obesity and diabetes mellitus. Breastfeeding also lowers maternal risk of breast and ovarian cancer, hypertension, and other metabolic diseases.^{3,45}
- Current computer system difficulties prevent the direct use of NC Medicaid diagnostic and cost data. However the Agency for Healthcare Research and Quality has done studies on the benefits of breastfeeding,⁶ and the results of these studies can be extrapolated to the NC Medicaid population.
- Such an extrapolation indicates that, if the NC Medicaid breastfeeding rates rise to the Healthy People 2020 goals, 14-18 infant deaths would be prevented.⁷
- In addition to the inestimable benefit of averting infant deaths, significant cost savings to NC Medicaid would be attained. In fact, the case for covering lactation support services can be made in the savings from just three common and sometimes lethal infections in infancy: Lower Respiratory Tract Infections; Gastroenteritis; and Necrotizing Enterocolitis.

Condition	NC Medicaid Cases Averted	Estimated Cost Savings
Lower Respiratory Tract Infections	500	\$2.5 Million
Gastroenteritis	6,000	\$2.5 Million
Necrotizing Enterocolitis	10	\$2.1 Million
Total Cost		\$7.1 Million

Cost Calculation

North Carolina Births	120,000
Births Covered by Medicaid	61,200 (51%)
# of Women on Medicaid initiating breastfeeding (Assuming 80% initiation goal is met)	48,960
Estimated Number of women who need lactation support visit (Estimated at 75%)	36,720
Estimated number of lactation support visits (1.3 visits)	47,736
Estimate cost of visits (assuming \$100/visit)	\$4.77 Million

- Preliminary cost / benefits analysis indicates that coverage of lactation support services by NC Medicaid would not exceed \$4.77 million annually.
- Extrapolation from national research studies indicates that 14-18 infant deaths would be prevented and more than \$7 million in Medicaid expenditures would be saved.
- Given this cost savings, providing reimbursement for appropriate medical lactation care services for North Carolina Medicaid infants is a financially sound investment.

References

1. Li R, Fein SB, Chen J, Grummer-Strawn LM. Why mothers stop breastfeeding: Mothers' self-reported reasons for stopping during the first year. *Pediatrics*. 2008;122 Suppl 2:S69-76.
2. U. S. Department of Health and Human Services. The surgeon general's call to action to support breastfeeding. . 2011:1-85.
3. Stuebe AM, Schwarz EB, Grewen K, et al. Duration of lactation and incidence of maternal hypertension: A longitudinal cohort study. *Am J Epidemiol*. 2011;174(10):1147-1158. doi: 10.1093/aje/kwr227; 10.1093/aje/kwr227.
4. Stuebe A. The risks of not breastfeeding for mothers and infants. *Rev Obstet Gynecol 2009 Fall*; 2(4): 222–231. 2009;2(4):222-231.
5. Bartick M, Stuebe A, Shealy KR, Walker M, Grummer-Strawn LM. Closing the quality gap: Promoting evidence-based breastfeeding care in the hospital. *Pediatrics*. 2009;124(4):e793-802.
6. Ip S, Chung M, Raman G, Trikalinos TA, Lau J. **A summary of the agency for healthcare research and quality's evidence report on breastfeeding in developed countries.** *Breastfeed Med 2009 Oct*;4 Suppl 1:S17-30. 2009;Suppl 1:s17-s30.
7. Bartick M, Reinhold A. The burden of suboptimal breastfeeding in the united states: A pediatric cost analysis. *Pediatrics*. 2010;125(5):e1048-56. doi: 10.1542/peds.2009-1616; 10.1542/peds.2009-1616.



July 17, 2013

Dear colleagues:

The [Patient Protection and Affordable Care Act](#) (ACA) – the health insurance reform legislation passed by Congress and signed into law by President Obama on March 23, 2010 – helps make preventive health affordable and accessible for all Americans by requiring health plans to cover preventive services and eliminating cost sharing for those services. Preventive services that have strong scientific evidence of their health benefits must be covered, and plans can no longer charge a patient a copayment, coinsurance, or deductible for these services when delivered by a network provider.

The United States Breastfeeding Committee (USBC) is pleased that the ACA requires coverage of [preventive health services for women](#), including “breastfeeding support, supplies, and counseling,” further defined as “comprehensive lactation support and counseling, by a trained provider during pregnancy and/or in the postpartum period, and costs for renting breastfeeding equipment.” These preventive services must be covered in conjunction with each birth, beginning in the first plan year (in the individual market, policy year) that begins on or after August 1, 2012.¹

Following the launch of this new coverage requirement, the USBC began to hear from member organizations, state breastfeeding coalitions, and the public of a pressing need for greater clarity on the definition of “breastfeeding support, supplies, and counseling.” In the rapidly-changing landscape of insurance requirements under the ACA, few guidelines or recommendations existed as to who may provide and be reimbursed for lactation care, and what kinds of equipment should be covered for breastfeeding families. As a result, inconsistent interpretations have emerged among insurers and health care providers, and families seeking breastfeeding support have experienced widespread confusion and disrupted continuity of care.

As a multi-sectoral, nonprofit coalition comprised of federal government agencies, non-governmental organizations, and health professional associations, the USBC serves as a national collective voice for supporting breastfeeding as a public health imperative. As such, the USBC was asked by member and partner organizations, state breastfeeding coalitions, and federal public health administrators² to prepare resources that could be made widely available to those seeking guidance for consistent payer coverage of breastfeeding services and equipment. The aim was to provide clear recommendations for federal and state agencies (that oversee delivery of health care services to eligible low-income families through Medicaid or similar programs), and for private insurers (that reimburse health care services through plans offered by employers or purchased by individuals).

In response to this need, the *Model Policy: Payer Coverage of Breastfeeding Counseling Services, Pumps, and Supplies* was created to **identify best practices for payers that appropriately meet the requirements of the ACA and ensure adequate delivery of breastfeeding support, using the language, format, and terminology applicable for expedient, meaningful use by this audience.** The *Model Policy* is intended to guide payers in determining insurance coverage. As such, it focuses on those health care professionals who are customarily credentialed by insurers³ to provide health care services in today’s insurance marketplace. This policy does not address community-based counseling and support options, offered by organizations and individuals in settings not currently covered by insurance.

Achieving equity in breastfeeding support services and access to care for all is of particular interest and concern to the USBC and breastfeeding advocates across the Nation. The USBC supports [*The Surgeon General's Call to Action to Support Breastfeeding*](#), in which Surgeon General Regina M. Benjamin emphasizes that “Everyone can help make breastfeeding easier.” Not all families will require skilled lactation care from a health care professional. Mother-to-mother support and peer counseling options are very effective means of providing support in a socially and culturally appropriate context. Breastfeeding peer counseling programs that support women enrolled in or eligible for WIC have been found to be effective at both agency and individual levels in improving breastfeeding rates.

As the health care landscape continues to evolve, innovative models are emerging for tiered care and/or partnerships between providers, so that all mothers have access to appropriate support to address common breastfeeding challenges, and receive referral to professional support in the event of complications. 'Incident-to' billing guidelines allow physicians to utilize non-licensed providers within their practices to assist with patient services. Provided that these guidelines are followed appropriately, non-licensed providers may be used to follow up with patients once the initial visit has been established with the physician.⁴

The USBC is very interested in exploring and advancing these models and identifying best practices for their promotion, adoption, and spread: feedback and examples from the community are welcome. The USBC is actively seeking funding to undertake such an effort collectively, especially working with health professional associations to encourage collaboration between providers to extend the reach of reimbursement, while still meeting payers' requirements for provider credentialing.

The *Model Policy* was drafted in a collaborative effort with medical, legal, regulatory, lactation, public health, and insurance experts, from the USBC Advocacy Committee and Board of Directors, and partner organizations the National Breastfeeding Center and California WIC Association. The National Women's Law Center also provided input. In accordance with USBC publications procedures, a draft was made available to all USBC member representatives, soliciting comments and feedback. In-depth comments from nine organizations were compiled, reviewed, and considered for incorporation.

The revised draft was reviewed by USBC legal counsel, and then the USBC Board reviewed and voted to publish it jointly with the National Breastfeeding Center, on July 8, 2013. The USBC Board appreciates the collective effort of the individuals and organizations whose insight, expertise, and perspective have allowed provision of this important resource to all those supporting breastfeeding families.

Sincerely,

The USBC Board of Directors

¹ For new plans created after August 1, 2012, and for non-grandfathered plans effective August 1, 2012.

² A meeting in March 2013 between USBC and California WIC Association representatives and Mayra Alvarez, Director of Public Health Policy, Office of Health Reform, U.S. Department of Health and Human Services (DHHS), confirmed that the period for additions/changes to the federal sub-regulatory guidance had passed. DHHS suggested that documents made available from the USBC website, as the national, nonprofit breastfeeding coalition, would be a valuable resource for the public and private sectors.

³ Insurance Credentialing (Provider Credentialing) is the process of becoming contracted with insurance companies so that the provider can participate in insurance company networks.

⁴ For information on 'incident-to' billing guidelines, see www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/downloads/se0441.pdf.



Model Policy

Payer Coverage of Breastfeeding Support and Counseling Services, Pumps and Supplies

© 2013 by the United States Breastfeeding Committee and National Breastfeeding Center. Cite as: United States Breastfeeding Committee, National Breastfeeding Center. *Model Policy: Payer Coverage of Breastfeeding Support and Counseling Services, Pumps and Supplies*. Washington, DC: United States Breastfeeding Committee and National Breastfeeding Center; 2013.

Background

Scientific evidence identifies breastfeeding as the optimal method of infant feeding. Support for breastfeeding, including lactation counseling services and breastfeeding pumps and supplies, ensures the opportunity for enhanced health for women and their children. Identified as the most compelling preventive modality available to health care providers, breastfeeding reduces the risk of several common causes of infant morbidity. Infants who are exclusively breastfed in the first six months of life (human milk only, along with needed medicine or vitamins), with complementary foods added thereafter while breastfeeding continues, demonstrate improved health and wellness as compared to their formula-fed peers. Most improved health outcomes manifest in a dose-response relationship: the greater the duration and/or exclusivity of breastfeeding, the greater the protective effect.

Breastfeeding's unique protective effect decreases the incidence of middle ear disease, diarrhea, and respiratory illnesses in the breastfed infant. Most research indicates the incidence of allergic disease, asthma, and obesity is less in infants who are breastfed. Significantly lower rates of necrotizing enterocolitis, a life-threatening bowel disorder related to prematurity, and Sudden Infant Death Syndrome (SIDS) occur among infants who are breastfed. The risk of developing chronic conditions such as childhood inflammatory bowel disease, lymphoma, leukemia, and insulin-dependent diabetes is significantly lower for breastfed infants.

Studies have shown that the benefits of breastfeeding to the mother include a reduced incidence of pre-menopausal breast cancer and pre-menopausal ovarian cancer. Newer evidence points to a reduced risk of developing Type 2 diabetes and a reduction in maternal cardiovascular disease.

The economic benefits to the family, health care system and workplace are widely published in the literature. If 90% of U.S. families could comply with the universal medical recommendation to breastfeed exclusively for six months, the United States is projected to save \$13 billion per year, including the cost of an excess 911 deaths, nearly all of which would be in infants (\$10.5 billion and 741

deaths at 80% compliance)¹. A similar study estimates the maternal health burden from current breastfeeding rates both in terms of premature death as well as economic costs, showing that suboptimal breastfeeding incurs a total of \$17.4 billion in cost to society resulting from premature death, \$733.7 million in direct costs, and \$126.1 million in indirect morbidity costs².

Current U.S. breastfeeding rates are suboptimal and result in significant excess costs and preventable infant deaths. Implementation of strategies to promote longer breastfeeding duration and exclusivity are needed to reduce health care costs and improve population health. For example, the Healthy People³ 2020 goals for breastfeeding include: breastfeeding initiation at 81.9%, and duration of any breastfeeding for six months at 60.6% and for one year at 34.1%. The most recent data from the U.S. Centers for Disease Control and Prevention (2009 births) identifies national breastfeeding initiation rates at 76.9%, duration of any breastfeeding for six months at 47.2% and for one year at 25.5%. There is clearly a significant gap to be bridged to meet the nation's public health goals for breastfeeding.

Professional health care organizations universally and actively encourage the support necessary to promote successful establishment and maintenance of breastfeeding. These organizations include (but the list is not limited to) the American Academy of Pediatrics; the American College of Obstetricians and Gynecologists; the American Academy of Family Physicians; the Academy of Breastfeeding Medicine; the American Hospital Association; the Association of Women's Health, Obstetric and Neonatal Nurses; the National Association of Pediatric Nurse Practitioners; the Academy of Nutrition and Dietetics; the American Public Health Association; the International Lactation Consultant Association; the United States Lactation Consultant Association; The Joint Commission; and the Human Milk Banking Association of North America. The United States Breastfeeding Committee (USBC), a multi-sectoral, nonprofit coalition comprised of federal government agencies, non-governmental organizations, and health professional associations, serves as a national collective voice for supporting breastfeeding as a public health imperative.

¹ Bartick M, Reinhold A. The burden of suboptimal breastfeeding in the United States: a pediatric cost analysis. *Pediatrics*. 2010;125(5):e1048-e1056.

² Bartick MC, Stuebe AM, Schwarz EB, Luongo C, Reinhold A, Foster EM. Cost analysis of maternal disease associated with suboptimal breastfeeding. *Obstet Gynecol*. 2013;122(1):111-119.

³ Healthy People provides a framework for health promotion and disease prevention for the nation and is designed to identify the most significant threats to public health and establish national goals to help reduce these threats.

Federal government agencies continue to offer significant data and documentation recognizing breastfeeding as a public health issue, supporting programs designed to increase breastfeeding rates. These organizations include (but the list is not limited to) Department of Health and Human Services: Centers for Disease Control and Prevention, Indian Health Service, Agency for Healthcare Research and Quality, Health Resources and Services Administration / Maternal and Child Health Bureau, Office on Women's Health, and Food and Drug Administration; and within the Department of Agriculture: the Food and Nutrition Service/Women, Infants and Children Program.

In addition, in January 2011, Surgeon General Regina M. Benjamin released *The Surgeon General's Call to Action to Support Breastfeeding*. This unprecedented document from the nation's highest medical source calls on health care providers, employers, insurers, policymakers, researchers, and the community at large to take 20 concrete action steps to support mothers in reaching their personal breastfeeding goals.

Policy

Health plan coverage guidelines developed by the Institute of Medicine, authorized under provisions of the Patient Protection and Affordable Care Act, and released by the U.S. Department of Health and Human Services (DHHS) require health insurance plans to cover certain women's preventive services including "breastfeeding support, supplies, and counseling." Insurers are required⁴ to provide coverage consistent with these guidelines, from the first plan year beginning August 1, 2012. **This model policy identifies best practices for supporting coverage of services, pumps and supplies appropriately.**

Breastfeeding Support and Counseling Services

In order to adequately support the intent of the law to ensure that women are adequately supported in their efforts to breastfeed, breastfeeding counseling services (also known as lactation consultation services) are a covered benefit during the prenatal and postpartum periods. Counseling services include

⁴ For new plans created after August 1, 2012, and for non-grandfathered plans effective August 1, 2012.

prenatal and postpartum breastfeeding education and support groups, as well as individual consultations to address specific problems, and are considered covered services. These include inpatient counseling and up to a total of six (6) outpatient lactation consultations. Additional consultations may be covered under medically necessary circumstances identified by and/or prescribed by a recognized provider.

This policy offers guidelines for coverage of health care providers who are professionally licensed, or, consistent with insurance companies' credentialing requirements, who have individual certification awarded by an independently-accredited program that measures assessment of predetermined standards for knowledge, skills, or competencies in a health-related profession, substantially equal to those articulated by the National Commission for Certifying Agencies (NCCA) and the Institute for Credentialing Excellence (ICE). The ability to counsel the breastfeeding mother and infant lies within the professional scope and practice of approved lactation care providers. International Board Certified Lactation Consultants (IBCLCs) are an example of non-licensed approved lactation care providers with certification meeting recognized credentialing standards.

Approved lactation care providers may vary in their training, licensure, certification, level of care, and ability to deliver care. Many disciplines offer lactation care to the mother and infant, such as nurses, advanced practice nurses, physician assistants, physicians, registered dietitians, peer lactation counselors, and lactation educators. The role of the IBCLC and other approved lactation care providers includes educating families, health professionals, and policymakers about the value of breastfeeding as a global public health imperative. These providers are trained to observe breastfeeding, assess, counsel, identify lactation problems, and specifically create a plan of care for the patients. This plan is communicated to other health care professionals and primary care providers based on an observed feeding assessment and follow-up care in order to help the mother and infant attain successful breastfeeding practices.

Breastfeeding Pumps and Supplies

Breast pumps and supplies are considered a covered benefit during the first year of the child's life, when provided from a durable medical equipment provider, an approved lactation care provider, a pump manufacturer or its representative, or an in-network provider, at the time of service whenever possible, or within 24 hours of notification of need. Shipping of the medical equipment is also covered.

Electric pump *rentals* ('hospital-grade' or 'rental-grade'⁵) are covered on a medically necessary⁶ basis only, such as to support initiation of lactation for mothers and infants who are separated due to illness, or who are unable to feed directly from the breast due to maternal or infant medical complications, congenital anomalies, prematurity, induced lactation, relactation, adoption, prematurity, or other medical conditions for mother or infant which preclude effective feeding at the breast.

Electric pump *purchases* (single- or multi-user) are a covered benefit for those mothers who need to maintain lactation when separated on a regular basis, or for a prolonged period of time, from their infant. One electric pump purchase will be covered every 36 months (or replaced during that period if the manufacturer's warranty is less than 36 months).

A breast pump kit is also covered, which allows expression from both breasts simultaneously (also known as 'double pumping'). This kit has tubing, valves, and other parts (specific to each manufacturer's requirements) as well as two flanges and two collection bottles for simultaneous pumping. Up to two double pump kits will be provided per pregnancy.

Other lactation equipment such as supplemental lactation aids, nipple shields, and breast shells may be covered when supplied during a lactation consultation and/or with the counsel of a permitted provider.

⁵ Sometimes called a "hospital-grade" or "rental-grade" pump, the preferred nomenclature identified by the U.S. Food and Drug Administration (FDA) is "multiple user" or "multi-user", to identify equipment intended for use by more than one person. See: <http://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/HomeHealthandConsumer/ConsumerProducts/BreastPumps/ucm061952.htm>, accessed June 27, 2013.

⁶ Mothers with a history of breast trauma, surgery, or delayed lactation benefit from the use of hospital-grade pumps, as documented in the medical literature.

Pasteurized donor human milk is considered life saving for fragile infants who are preterm or who are medically compromised, such as those diagnosed with formula intolerance, metabolic conditions, or genetic conditions requiring human milk. Maternal illness, maternal death, surgery, drug or medication use contraindicated for breastfeeding, and chronic conditions that are contraindicated for breastfeeding, are substantive reasons for use of pasteurized donor human milk. Pasteurized donor human milk from a screened donor provided by a human milk bank associated with the Human Milk Banking Association of North America (HMBANA),⁷ will be a covered benefit. Human milk purchases provided by non-HMBANA milk-sharing groups will not be covered. Prescription from a licensed provider identifying the medical condition/diagnosis and necessity for the infant to receive pasteurized donor human milk is required for this coverage. Prescriptions for banked donor human milk must be renewed every three months.

⁷ HMBANA: The Human Milk Banking Association of North America is a non-profit organization that provides guidelines for pasteurization of milk from carefully screened donors for fragile and sick infants improving their potential to survive and thrive.

Breastfeeding Support & Counseling Services Benefit Guidelines

	Coverage	Permitted Providers	Coverage Criteria	Member Costs, Exclusions, Maximums
Education & Support				
Prenatal/postpartum breastfeeding classes	Covered benefit for one class series in the prenatal period and one class series up to 12 months postpartum	Approved lactation care providers, ⁸ licensed providers (such as MD, APRN, DO, PA), contracted network providers, RN certified childbirth/breastfeeding educators, and Registered Dietitians (RD, RDN, LDN) ⁹	No referral or prior authorization required	No cost to member Prenatal and postpartum series up to a total of 18 class weeks
Breastfeeding support groups	Covered benefit for one monthly meeting up to a maximum of 12 meetings	Approved lactation care providers, licensed providers (such as MD, APRN, DO, PA), contracted network providers, RN certified childbirth/breastfeeding educators, and Registered Dietitians (RD, RDN, LDN)	No referral or prior authorization required	No cost to member One monthly meeting up to a total of 12 meetings
Lactation Counseling				
In-patient hospital	Covered benefit included as part of inpatient stay for facility employed providers; separately payable service (with place of service code) for non-employee in-network providers	Network hospitals, employed approved lactation care and/or licensed providers; in-network non-employee providers	No referral or prior authorization required	No cost to member
Outpatient hospital, home, provider office/clinic	Covered benefit for 6 outpatient lactation consultation visits per birth up to 2 hours per visit	Approved lactation care providers, licensed providers (such as MD, APRN, DO, PA); non-licensed providers, RN, and RD, RDN, LDN may bill as 'incident-to' physician services ¹⁰	No referral or prior authorization required; prior authorization required beyond 6 visits	No cost to member Up to 6 visits per birth, up to 2 hours per visit

⁸ Approved lactation care providers are those who, consistent with insurance companies' credentialing requirements, have individual certification awarded by an independently-accredited program that measures assessment of predetermined standards for knowledge, skills, or competencies in a health-related profession, substantially equal to those articulated by the National Commission for Certifying Agencies and the Institute for Credentialing Excellence.

⁹ RD (registered dietitian), RDN (registered dietitian and nutritionist), LDN (licensed dietitian and nutritionist)

¹⁰ For information on 'incident-to' billing guidelines, see www.cms.gov/Outreach-and-Education/Medicare-Learning-Network-MLN/MLNMattersArticles/downloads/se0441.pdf, accessed April 9, 2013.

Breastfeeding Pumps and Supplies Benefit Guidelines

	Coverage	Permitted Providers	Coverage Criteria	Member Costs, Exclusions, Maximums
Multi-user rental-grade electric breast pump ☐ Rental (CPT code: E0604)	Covered benefit per birth event, with prescription indicating medical necessity	Hospitals, DME approved vendors, approved lactation care providers, in-network providers, or direct order from manufacturers or their representatives ☐ To be provided at time of service or within 24 hours of notification of need ☐ Includes instructions on assembly, cleaning, and use	☐ Requires prescription by a licensed provider, specifying medical necessity and length of need, as advised by an approved lactation care provider	No cost to member Maximum benefit is \$1,000 for rental and kit, for up to one year
Electric breast pump (single- or multi-user)¹¹ ☐ Purchase (CPT code: E0603)	Covered benefit per birth event, with prescription	DME approved vendors, approved lactation care providers, in-network providers, or direct order from manufacturers or their representatives ☐ To be provided at time of service or within 24 hours of notification of need ☐ Includes instructions on assembly, cleaning, and use	☐ No 'medical necessity' requirement ☐ Requires prescription by a licensed provider, as advised by an approved lactation care provider ☐ Prescription should specify pump (make and model) specific to individual's needs	No cost to member Maximum benefit is \$1,000 for pump and accessories One breast pump purchase covered every 36 months, or replaced per pregnancy where manufacturer's warranty is less than 36 months

¹¹ Mothers that may experience limited access to electricity will be provided with a manual pump (that includes a separate collection container with a cylinder or lever to promote suction and release) upon request, in addition to the appropriate single-user or multi-user electric breast pump.

Pump kits	Covered benefit, no prescription needed	DME approved vendors, approved lactation care providers, in-network providers, or direct order from manufacturers or their representatives ☐ Double pump kits are provided with the rental or purchase of a single- or multi-user electric pump	☐ Initial kit included with pump ☐ Second kit covered with medical necessity ☐ Additional kit upon request per birth if within 36 months of pump purchase	No cost to member New kit covered with every birth event, up to a maximum benefit of \$150 per kit
Other equipment, supplemental lactation aid,¹² nipple shields,¹³ breast shells¹⁴	Covered benefit, no prescription needed	DME approved vendors, approved lactation care providers, in-network providers, or direct order from manufacturers or their representatives	☐ Requires coding V24.1	No cost to member Maximum benefit of \$250 per calendar year
Donor human milk^{7,15}	Covered benefit for identified at-risk infants through the first 12 months of life	Supplied through HMBANA-associated milk banks, or banks of equal or greater standard, only ⁷	☐ Requires prescription by a licensed provider; renewable every three months	No cost to member Maximum 12 month supply

¹² Supplemental Lactation Aid: Brand names include Supplemental Nursing System (SNS), starter SNS, Lact-Aid and others used to provide a supplement of additional milk while the infant is nursing at the breast. This supplement is delivered during the breastfeeding process through the use of a small tube attached to the breast and milk flowing from a container through the tube to the infant.

¹³ Nipple Shields: A nipple shield is made of thin flexible silicone and is used to aid in the breastfeeding latching process or when the infant and mother have difficulty breastfeeding due to sore or flat nipples. It is a reusable product with the mother only requiring a maximum of two shields.

¹⁴ Breast Shells: Breast shells are used to protect sore nipples during the healing process. They are worn inside the bra and removed prior to feeding.

¹⁵ Donor Human Milk: Action 12 of *The Surgeon General's Call to Action to Support Breastfeeding* states, "Identify and address obstacles to greater availability of safe banked donor milk for fragile infants."

A note regarding breast pumps¹⁶:

Breast pumps vary in terms of quality. *Appropriate* pumps will have these features: automatic electric breast pump with intermittent suction, 50-80 cycles per minute with adjustable vacuum ranging from 50 to 250 mm Hg. A pedal pump used by “foot” power may be substituted for electric pumps for mothers with compromised upper body arms and/or hand anomalies. Please see Appendix for details regarding Minimum Breast Pump Specifications.

Medical documentation and referral for breastfeeding infant with special needs:

Special needs children should be seen as soon as possible after delivery and be referred to an approved lactation care provider at discharge, in order to receive services no later than day 3-4 postpartum. Mothers with history of infertility, breast surgery, or special medical needs should see inpatient approved lactation care providers for an initial consultation within 24 hours of delivery, and after discharge by day 3-4 postpartum.

Nomenclature

Multi-User or ‘Rental-Grade’ Electric Pump: These breast pumps are utilized in the hospital setting and by rental after discharge. There is an implication that these types of breast pumps are intended for multiple, different users for mother-infant dyads facing significant medical challenges. As a result these breast pumps typically have enhanced features to meet the demands of these medical challenges. These pumps have also been referred to as ‘hospital grade’ in the consumer market. The designation of ‘hospital grade’ is not recognized by the FDA. This term originated due to the pumps having a three-pronged grounded plug. These heavy-duty electric breast pumps are rented and returned for reuse by many different women.

Multi-User Electric Pump: This is an FDA designation indicating that this type of breast pump prevents contamination with infectious particles and organisms between different users (cross contamination) due to the pump's filter/design. The multi-user designation does not specifically address where the pump should be

¹⁶ For definitive information on pump quality and efficacy, please refer to *Clinics in Human Lactation: Pumps & Pumping Protocols*. Walker, Marsha. Hale Publishing, 2012.

used (Neonatal Intensive Care Unit, postpartum unit, home, school, or business). Any pump that would be shared or used by multiple women should be a multi-user breast pump to prevent potential infection between users. Different women using the same multi-user pump should have their own individual, personal-use kit (that connects the pump to both breasts simultaneously facilitating milk extraction and collection) with tubing/flanges/bottles that attach to the pump. Multi-user designed pumps are also used by single-users. The multi-user designation does not address the pump motor or any other pump specifications. Some multi-user pumps have 36-month warranties and may be preferable to 12-month warranted pumps.

Single User Electric Pump: This is not an officially designated description. The 'single-user' descriptor is associated with breast pumps that are only to be used by one person.

Double (sided) Electric Pump: This refers to a breast pump that can simultaneously express both breasts. These pumps utilize appropriate kits for the pump, including flanges, bottles, and tubing that connect the pump to both breasts simultaneously to facilitate milk extraction and collection. Double electric pumps are recognized as the standard pump in both the inpatient and outpatient setting. Single (sided) electric pumps are inferior to double (sided) electric pumps.

Electric Pump: These utilize electricity for power from standard wall or car outlets.

Battery and Manual (hand) Pumps: These pumps utilize either battery power or are powered by hand and are inferior to electric pumps.

Coding Guidelines

ICD9 / DIAGNOSIS CODES	
INFANT CODES	
Infant Feeding Problems	
779.31	Feeding problems, newborn
779.32	Bilious vomiting in newborn
779.33	Other vomiting in newborn
783.3	Feeding problems >28 days
787.03	Vomiting, infant >28 days
Gestation and Congenital Disorders	
765.1	Prematurity
765.22-765.29	Specified weeks of gestation/prematurity
740-759	Congenital anomalies
Jaundice	
774.39	Breastmilk jaundice
774.6	Neonatal jaundice, unspecified
774.2	Preterm jaundice
Weight and Hydration	
775.5	Dehydration, neonatal
779.34	Failure to thrive, newborn
783.21	Weight loss
783.41	Failure to thrive, >28 days
783.1	Abnormal weight gain (poor weight gain)
Infant Distress	

780.91-780.92	Fussy infant/baby/excessive crying
789.7	Infantile colic or intestinal distress
GI Issues	
530.81	Gastroesophageal reflux
787.7	Abnormal stools
787.91	Diarrhea
Mouth	
112.0	Oral thrush
749.00-749.25	Cleft Palate and Cleft Lip
750.0-750.19	Ankyloglossia and other anomalies of the tongue
750.21-750.29	High Arched Palate and other anomalies of mouth and pharynx
25.92	Lingual Frenectomy
41010	Incision tongue fold
Airway	
748.3	Tracheomalacia
519.19	Laryngomalacia
MATERNAL CODES	
675.14	Abscess, breast/Mastitis, infective
675.24	Blocked milk duct/Mastitis, interstitial
676.24	Breast engorgement, ductal
782.0	Burning pains, hyperesthesia
757.6	Ectopic or axillary breast tissue/Other specified nipple/breast anomaly

675.84	Other specified nipple/breast anomalies
676.84	Galactoceles
675.04	Nipple Infection
676.14	Nipple cracks/fissures
676.34	Sore nipples
676.04	Retracted Nipple
684	Staph/ impetigo
112.89	Candidiasis, nipple/breast
Constitutional	
780.55	Disruption of sleep cycle
780.79	Fatigue
684.44	Mental disorders postpartum
642.0-642.7	Hypertension complications during pregnancy and childbirth
660-669	Complications Occurring during Labor and Delivery
670-676	Complications of the Puerperium
678-679	Other maternal and fetal complications
Lactation	
676.44	Agalactia, failure to lactate
676.84	Delayed lactation/specified disorders of lactation
676.54	Lactation suppressed
V24.1	Supervision of lactation
V24.0-V24.2	Postpartum care and examination
V67.59	Other specified follow-up exam when original reason for visit has resolved
Pregnancy	

651.00	Twin Pregnancy, unspecified as to episode of care
651.10	Triplet Pregnancy, unspecified as to episode of care
651.13	Triplet Pregnancy, antepartum condition or complication
651.20	Quadruplet Pregnancy, unspecified as to episode of care
651.90	Unspecified multiple gestation, unspecified as to episode of care
HCPCS and CPT CODES	
96150	Initial assessment by lactation consultant of previously diagnosed problem
96151	Reassessment to measure goals formulated in initial assessment, and to modify plans as needed
99212	Problem Focused
99213	Expanded Problem Focused
99214	Detailed Problem Focused
99215	Comprehensive Exam
99201-99205	Straight forward- Comprehensive High Medical Decision Making
99241-99245	Consultation codes: straight forward- Comprehensive
99401-99404	Preventative Counseling Codes (15 minute increments 99401=15 minutes, 99404=60 minutes)
99411-99412	Group Education
99078	Physician educational services rendered to patients in a group setting (e.g., prenatal, obesity, or diabetic instructions)
S9443	Lactation Classes, Non-Physician Provider, Per Session

98960	Education And Training For Patient Self-Management By A Qualified, Non-physician Health Care Professional Using A Standardized Curriculum, Face-To-Face With The Patient (Could Include Caregiver/Family) Each 30 Minutes; Individual Patient
98961	Education and Training Each 30 Minutes; 2-4 Patients
98962	Education and Training for each 30 minutes; 5-8 patients
Modifier 33	Preventive Service
Pumps, Supplies and Equipment	
A4281-A4286	Breast pump supplies
E0603	Breast pump Electric (AC and /or DC) single- or multi-user
E0604	Breast pump Electric (AC and /or DC) 'rental grade'

Bibliography

Reference Sources

Academy of Breastfeeding Medicine Clinical Protocol Committee. ABM clinical protocol #2 (2007 revision): guidelines for hospital discharge of the breastfeeding term newborn and mother: "the going home protocol". *Breastfeed Med*. 2007;2(3):158-165.

Academy of Breastfeeding Medicine Protocol Committee. ABM clinical protocol #3: hospital guidelines for the use of supplementary feedings in the healthy term breastfed neonate, revised 2009. *Breastfeed Med*. 2009;4(3):175-182.

American Academy of Family Physicians. Family physicians supporting breastfeeding (position paper). <http://www.aafp.org/about/policies/all/breastfeeding-support.html>. Published 2001. Updated 2008. Accessed July 15, 2013.

American Academy of Pediatrics. Human milk. In: Pickering LK, Baker CJ, Long SS, McMillan JA, eds. *Red Book: 2006 Report of the Committee on Infectious Diseases*. 27th ed. Elk Grove Village, IL: American Academy of Pediatrics; 2006: 123-130.

American Academy of Pediatrics Section on Breastfeeding. Breastfeeding and the use of human milk (policy statement). *Pediatrics*; originally published online February 27, 2012; DOI: 10.1542/peds.2011-3552.

American Academy of Pediatrics Task Force on Sudden Infant Death Syndrome. The changing concept of Sudden Infant Death Syndrome: diagnostic coding shifts, controversies regarding the sleeping environment, and new variables to consider in reducing risk. *Pediatrics*. 2005;116(5):1245-1255.

Angeletti MA. Breastfeeding mothers returning to work: possibilities for information, anticipatory guidance and support from U.S. health care professionals. *J Hum Lact*. 2009;25(2):226-232.

Ball TM, Wright AL. Health care costs of formula-feeding in the first year of life. *Pediatrics*. 1999;103(4, pt 2):870-876.

Bartick M, Reinhold A. The burden of suboptimal breastfeeding in the United States: a pediatric cost analysis. *Pediatrics*. 2010;125(5):e1048-56.

Bartick MC, Stuebe AM, Schwarz EB, Luongo C, Reinhold A, Foster EM. Cost analysis of maternal disease associated with suboptimal breastfeeding. *Obstet Gynecol*. 2013;122(1):111-119.

Becker GE, Cooney F, Smith HA. Methods of milk expression for lactating women. *Cochrane Database Syst Rev*. 2011;(12):CD006170.

Bramson L, Lee JW, Moore E, et al. Effect of early skin-to-skin mother-infant contact during the first 3 hours following birth on exclusive breastfeeding during the maternity hospital stay. *J Hum Lact*. 2010;26(2):130-137.

Britton C, McCormick FM, Renfrew MJ, Wade A, King SE. Support for breastfeeding mothers. *Cochrane Database Syst Rev*. 2007;(1):CD001141.

California WIC Association. Opportunities for Nutrition and Breastfeeding Interventions Under Health Care Reform. http://www.calwic.org/storage/documents/reports/CWA-Lactation_and_Nutrition_Counseling_Under_Health_Reform_May_2012.pdf. Published May 2012. Accessed June 27, 2013.

California WIC Association. Ramping Up for Reform: Quality Breastfeeding Support in Preventive Care. http://www.calwic.org/storage/documents/bf/2012/Ramping_Up_for_Reform-WIC_Breastfeeding_Toolkit_2012.pdf. Published February 2012. Accessed June 27, 2013.

Castrucci BC, Hoover KL, Lim S, Maus KC. A comparison of breastfeeding rates in an urban birth cohort among women delivering infants at hospitals that employ and do not employ lactation consultants. *J Public Health Manag Pract*. 2006;12(6):578-585.

Centers for Disease Control and Prevention, Division of Nutrition, Physical Activity Obesity, National Center for Chronic Disease Prevention Health Promotion. Breastfeeding among U.S. children born 2000-2009, CDC National Immunization Survey. Centers for Disease Control and Prevention Web site. www.cdc.gov/breastfeeding/data/NIS_data/index.htm. Updated August 1, 2012. Accessed June 27, 2013.

Committee on Health Care for Underserved Women, American College of Obstetricians and Gynecologists. ACOG committee opinion no. 361: breastfeeding: maternal and infant aspects. *Obstet Gynecol*. 2007;109(2, pt 1):479-480.

Cheng PG, Johnson LW, Rosenthal MS. Sources of education about breastfeeding and breast pump use: what effect do they have on breastfeeding duration? An analysis of the Infant Feeding Practices Survey II. *Matern Child Health J*. 2012;16(7):1421-1430.

Chung M, Raman G, Trikalinos T, Lau J, Ip S. Interventions in primary care to promote breastfeeding: an evidence review for the U.S. Preventive Services Task Force. *Ann Intern Med*. 2008;149(8):565-582.

Declercq E, Labbok MH, Sakala C, O'Hara M. Hospital practices and women's likelihood of fulfilling their intention to exclusively breastfeed. *Am J Public Health*. 2009;99(5):929-935.

Dennis CL. Identifying predictors of breastfeeding self-efficacy in the immediate postpartum period. *Res Nurs Health*. 2006;29(4):256-268.

Dewey KG, Nommsen-Rivers LA, Heinig MJ, Cohen RJ. Risk factors for suboptimal infant breastfeeding behavior, delayed onset of lactation, and excess neonatal weight loss. *Pediatrics*. 2003;112(3, pt 1):607-619.

- DiGirolamo AM, Grummer-Strawn LM, Fein SB. Effect of maternity-care practices on breastfeeding. *Pediatrics*. 2008;122(suppl 2):S43-S49.
- Donnelly A, Snowden HM, Renfrew MJ, Woolridge MW. Commercial hospital discharge packs for breastfeeding women. *Cochrane Database Syst Rev*. 2000;(2):CD002075.
- Dyson L, McCormick F, Renfrew MJ. Interventions for promoting the initiation of breastfeeding. *Cochrane Database Syst Rev*. 2005;(2):CD001688.
- Feldman-Winter L, Barone L, Milcarek B, et al. Residency curriculum improves breastfeeding care. *Pediatrics*. 2010;126(2):289-297.
- Gagnon AJ, Leduc G, Waghorn K, Yang H, Platt RW. In-hospital formula supplementation of healthy breastfeeding newborns. *J Hum Lact*. 2005;21(4):397-405.
- Gonzalez KA, Meinen-Derr J, Burke BL, et al. Evaluation of a lactation support service in a children's hospital neonatal intensive care unit. *J Hum Lact*. 2003;19(3):286-292.
- Grossman X, Chaudhuri J, Feldman-Winter L, et al. Hospital Education in Lactation Practices (Project HELP): does clinician education affect breastfeeding initiation and exclusivity in the hospital? *Birth*. 2009;36(1):54-59.
- Guisse JM, Palda V, Westhoff C, Chan BK, Helfand M, Lieu TA; U.S. Preventive Services Task Force. The effectiveness of primary care-based interventions to promote breastfeeding: systematic evidence review and meta-analysis for the US Preventive Services Task Force. *Ann Fam Med*. 2003;1(2):70-78.
- Hannula L, Kaunonen M, Tarkka MT. A systematic review of professional support interventions for breastfeeding. *J Clin Nurs*. 2008;17(9):1132-1143.
- Hayes DK, Prince CB, Espinueva V, Fuddy LJ, Li R, Grummer-Strawn LM. Comparison of manual and electric breast pumps among WIC women returning to work or school in Hawaii. *Breastfeed Med*. 2008;3(1):3-10.
- Horta BL, Bahl R, Martines JC, Victora CG; World Health Organization. *Evidence on the Long-Term Effects of Breastfeeding: Systematic Reviews and Meta-Analysis*. Geneva, Switzerland: World Health Organization; 2007.
- Howard C, Howard F, Lawrence R, Andresen E, DeBlieck E, Weitzman M. Office prenatal formula advertising and its effect on breast-feeding patterns. *Obstet Gynecol*. 2000;95(2):296-303.
- Ip S, Chung M, Raman G, et al. *Breastfeeding and Maternal and Infant Health Outcomes in Developed Countries*. Rockville, MD: Agency for Healthcare Research and Quality; 2007. Evidence Report/Technology Assessment No. 153.
- Johnston ML, Esposito N. Barriers and facilitators for breastfeeding among working women in the United States. *J Obstet Gynecol Neonatal Nurs*. 2007;36(1):9-20.

- Keister D, Roberts KT, Werner SL. Strategies for breastfeeding success. *Am Fam Physician*. 2008;78(2):225-232.
- Kramer MS, Kakuma R. Optimal duration of exclusive breastfeeding. *Cochrane Database Syst Rev*. 2012;(8):CD003517.
- Mannel R, Mannel RS. Staffing for hospital lactation programs: recommendations from a tertiary care teaching hospital. *J Hum Lact*. 2006;22(4):409-417.
- MCH Leadership Competencies Workshop. *Maternal and Child Health Leadership Competencies*. Version 3.0. Rockville, MD: U.S. Department of Health and Human Services, Health Resources Services Administration, Maternal and Child Health Bureau; 2009.
- McNiel ME, Labbok MH, Abrahams SW. What are the risks associated with formula feeding? A re-analysis and review. *Breastfeed Rev*. 2010;18(2):25-32.
- Mulder PJ, Johnson TS. The Beginning Breastfeeding Survey: measuring mothers' perceptions of breastfeeding effectiveness during the postpartum hospitalization. *Res Nurs Health*. 2010;33(4):329-344.
- Murray EK, Ricketts S, Dellaport J. Hospital practices that increase breastfeeding duration: results from a population-based study. *Birth*. 2007;34(3):202-211.
- Naylor AJ. Baby-Friendly Hospital Initiative. Protecting, promoting and supporting breastfeeding in the twenty-first century. *Pediatr Clin North Am*. 2001;48(2):475-483.
- Neifert MR. Prevention of breastfeeding tragedies. *Pediatr Clin North Am*. 2001;48(2):273-297.
- Noble LM, Noble A, Hand IL. Cultural competence of healthcare professionals caring for breastfeeding mothers in urban areas. *Breastfeed Med*. 2009;4(4):221-224.
- Ohya M, Watabe H, Hayasaka Y. Manual expression and electric breast pumping in the first 48 h after delivery. *Pediatr Int*. 2010;52(1):39-43.
- Pak-Gorstein S, Haq A, Graham EA. Cultural influences on infant feeding practices. *Pediatr Rev*. 2009;30(3):e11-e21.
- Pérez-Escamilla R. Evidence based breast-feeding promotion: the Baby-Friendly Hospital Initiative. *J Nutr*. 2007;137(2):484-487.
- Philipp BL; Academy of Breastfeeding Medicine Protocol Committee. ABM clinical protocol #7: model breastfeeding policy (revision 2010). *Breastfed Med*. 2010;5(4):173-177.
- Philipp BL, McMahon MJ, Davies S, Santos T, Jean-Marie S. Breastfeeding information in nursing textbooks needs improvement. *J Hum Lact*. 2007;23(4):345-349.

Philipp BL, Merewood A, Gerendas EJ, Bauchner H. Breastfeeding information in pediatric textbooks needs improvement. *J Hum Lact*. 2004;20(2):206-210.

Quigley MA, Henderson G, Anthony MY, McGuire W. Formula milk versus donor breast milk for feeding preterm or low birth weight infants. *Cochrane Database Syst Rev*. 2007;(4):CD002971.

Renfrew MJ, Dyson L, McCormick F, et al. Breastfeeding promotion for infants in neonatal units: a systematic review. *Child Care Health Dev*. 2010;36(2):165-178.

Rosen IM, Krueger MV, Carney LM, Graham JA. Prenatal breastfeeding education and breastfeeding outcomes. *MCN Am J Matern Child Nurs*. 2008;33(5):315-319.

Rosenberg KD, Stull JD, Adler MR, Kasehagen LJ, Crivelli-Kovach A. Impact of hospital policies on breastfeeding outcomes. *Breastfeed Med*. 2008;3(2):110-116.

Schwarz EB, Ray RM, Stuebe AM, et al. Duration of lactation and risk factors for maternal cardiovascular disease. *Obstet Gynecol*. 2009;113(5):974-982.

Shaikh U, Smillie CM. Physician-led outpatient breastfeeding medicine clinics in the United States. *Breastfeed Med*. 2008;3(1):28-33.

Shealy KR, Li R, Benton-Davis S, Grummer-Strawn LM. *The CDC Guide to Breastfeeding Interventions*. Atlanta, GA: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2005.

Slavit W, ed. *Investing in Workplace Breastfeeding Programs and Policies: An Employer's Toolkit*. Washington, DC: Center for Prevention and Health Services, National Business Group on Health; 2009.

Spatz DL. Ten steps for promoting and protecting breastfeeding for vulnerable infants. *J Perinat Neonatal Nurs*. 2004;18(4):385-396.

Stuebe AM, Schwarz EB. The risks and benefits of infant feeding practices for women and their children. *J Perinatol*. 2010;30(3):155-162.

Szucs KA, Miracle DJ, Rosenman MB. Breastfeeding knowledge, attitudes, and practices among providers in a medical home. *Breastfeed Med*. 2009;4(1):31-42.

Taveras EM, Li R, Grummer-Strawn L, et al. Opinions and practices of clinicians associated with continuation of exclusive breastfeeding. *Pediatrics*. 2004;113(4):e283-e290.

U.S. Department of Health and Human Services. 2020 topics & objectives: Maternal, Infant, and Child Health. HealthyPeople.gov Web site. <http://healthypeople.gov/2020/default.aspx>. Updated March 8, 2013. Accessed March 13, 2013.

U.S. Department of Health and Human Services. *The Surgeon General's Call to Action to Support Breastfeeding*. Washington, DC: U.S. Department of Health and Human Services, Office of the Surgeon General; 2011.

United States Breastfeeding Committee. *Core Competencies in Breastfeeding Care and Services for All Health Professionals*. Rev ed. Washington, DC: United States Breastfeeding Committee; 2010.

Wagner J, Hanson C, Anderson Berry A. Donor human milk for premature infants: a review of current evidence. *Infant Child Adolesc Nutr*. 2013;5(2):71-77.

Walker M. *Clinics in Human Lactation: Pumps & Pumping Protocols*. Amarillo, TX: Hale Publishing; 2012.

Walker M. Conquering common breast-feeding problems. *J Perinat Neonatal Nurs*. 2008;22(4):267-274.

Wolynn T. Breastfeeding—so easy even a doctor can support it. *Breastfeed Med*. 2011;6(5):345-347.

Related Readings

Bergström A, Okong P, Ransjö-Arvidson AB. Immediate maternal thermal response to skin-to-skin care of newborn. *Acta Paediatr.* 2007;96(5):655-658.

Bhandari N, Bahl R, Mazumdar S, Martines J, Black RE, Bhan MK; Infant Feeding Study Group. Effect of community-based promotion of exclusive breastfeeding on diarrhoeal illness and growth: a cluster randomized controlled trial. *Lancet.* 2003;361(9367):1418-1423.

Brent NB, Redd B, Dworetz A, D'Amico F, Greenberg JJ. Breast-feeding in a low-income population. Program to increase incidence and duration. *Arch Pediatr Adolesc Med.* 1995;149(7):798-803.

Dykes F. The education of health practitioners supporting breastfeeding women: time for critical reflection. *Matern Child Nutr.* 2006;2(4):204-216.

Mittman BS. Creating the evidence base for quality improvement collaboratives. *Ann Intern Med.* 2004;140(11):897-901.

Moore ER, Anderson GC, Bergman N, Dowswell T. Early skin-to-skin contact for mothers and their healthy newborn infants. *Cochrane Database Syst Rev.* 2012;(5):CD003519.

Mori R, Khanna R, Pledge D, Nakayama T. Meta-analysis of physiological effects of skin-to-skin contact for newborns and mothers. *Pediatr Int.* 2010;52(2):161-170.

Rosenberg KD, Eastham CA, Kasehagen LJ, Sandoval AP. Marketing infant formula through hospitals: the impact of commercial hospital discharge packs on breastfeeding. *Am J Public Health.* 2008;98(2):290-295.

Rosswurm MA, Larrabee JH. A model for change to evidence-based practice. *Image J Nurs Sch.* 1999;31(4):317-322.

Spatz DL, Pugh LC; American Academy of Nursing Expert Panel on Breastfeeding. The integration of the use of human milk and breastfeeding in baccalaureate nursing curricula. *Nurs Outlook.* 2007;55(5):257-263.

Appendix: Minimum Breast Pump Specifications¹⁷

Minimum Specifications for Single- or Multi-User Double Electric Breast Pumps (E0603)

The electric personal use pump must:

- Be lightweight and portable. The total weight of furnished assembly should not exceed ten (10) pounds.
- Be packaged pre-assembled with all parts necessary for pumping. Assembly includes but not limited to pump motor unit, minimum 5' electric cord, and double pumping collection kit.
- Operate on a 110-volt household current and be UL listed.
- Have an adjustable suction pressure between 50 mm Hg and 250 mm Hg at the breast shield during use; a suction range just at the low or high end of the range is not acceptable.
- Have an automatic mechanism to prevent suction greater than 250 mm Hg when used according to manufacturer instructions to prevent nipple trauma.
- Have a mechanism for automatic release of suction for safety.
- Have variable/adjustable cycling not less than 30 cycles per minute; one fixed cycling time is not acceptable.
- Have single and double pumping capacity and capable of maintaining a consistent vacuum (no pressure change) as the collection container fills regardless of the container size and whether single or double pumping.
- Have double pumping capacity which is simultaneous, not alternating.
- Have a visible breast milk pathway and no milk is able to contact the internal pump motor unit parts at any time when the product is used per manufacturer instructions.
- Include breast flanges that are either adjustable/flexible or if rigid, come in at least two (2) sizes to accommodate different breast sizes with no sharp edges.
- Include a collection bottle of four (4) to six (6) ounces with a spill-proof cap and standard-size opening, and be bisphenol-A (BPA) and DHEP-free.
- Include a durable soft-sided carrying case with a storage compartment to hold pumping accessories and an insulated cooling compartment including freezer packs for storing expressed breast milk; this is recommended especially for women returning to work or school.
- Include a battery option and adapter that can be used as an alternate power source other than electric; this is recommended for flexibility of pumping.

¹⁷ Based on New York State Minimum Breast Pump Specifications for Medicaid Reimbursement, http://www.health.ny.gov/community/pregnancy/breastfeeding/medicaid_coverage/minimum_breast_pump_specifications.htm, accessed June 27, 2013.

Minimum Specifications for Single- or Multi-User Double Electric Pumping Kits

The kit must:

- Include breast flanges that are either adjustable/flexible or if rigid, come in at least two (2) sizes to accommodate different breast sizes with no sharp edges.
- Be packaged pre-assembled with all accessories necessary for pumping two breasts simultaneously or only one breast manually.
- Include at least two collection bottles of four (4) to six (6) ounces with a spill-proof cap and standard-sized opening, and be bisphenol-A (BPA) and DHEP-free.
- Contain collection bottle(s) and flanges made of medical grade quality to allow for repeated boiling and/or dishwasher cleaning, and which are scratch resistant and non-breakable.
- Have durable tubing designed for long-term pumping use.
- Design and materials of the furnished assembly shall allow viewing the breast milk pathway.
- Include a battery option and adapter that can be used as an alternate power source other than electric; this is recommended and may come as part of pump assembly or pumping kit.

Minimum Specifications for Rental-Grade Double Electric Breast Pump (E0604)

The electric rental-grade multi-user pump must:

- Must not exceed 12 pounds including carrying case.
- Operate on a 110-volt household current and be UL listed.
- Have a visible breast milk pathway and no milk is able to contact the internal pump motor unit parts at any time when the product is used per manufacturer instructions.
- Have an adjustable suction pressure between 30 mm Hg and 250 mm Hg at the breast shield during use; a suction range just at the low or high end of the range is not acceptable.
- Have an automatic mechanism to prevent suction greater than 250 mm Hg when used according to manufacturer instructions to prevent nipple trauma.
- Have a mechanism for automatic release of suction for safety.
- Have variable/adjustable cycling not less than 30 cycles per minute; one fixed cycling time is not acceptable.
- Have double pumping capacity, which is simultaneous, not alternating.
- Include a pumping kit for each personal user including durable tubing to connect to the pump and flanges, and have single and double pumping capacities.
- Include a carrying case made of durable, washable materials for the pump motor assembly and pump kit accessories; this is recommended if the pump needs to be portable.