Health Evidence Review Commission's Health Technology Assessment Subcommittee

November 25, 2013

1:00 PM

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

Section 1.0

Call to Order

AGENDA

HEALTH TECHNOLOGY ASSESSMENT SUBCOMMITTEE (HTAS)

Meridian Park Hospital Health
Education Center, Room 117B&C
19300 SW 65th Avenue, Tualatin, OR 97062
November 25, 2013
1:00 pm - 4:00 pm

All agenda items are subject to change and times listed are approximate

Public comment on listed topics will be taken at the time that topic is discussed and may be limited depending on the number of individuals providing testimony

#	Time	Item	Presenter	Action Item
1	1:00 PM	Call to Order	Alissa Craft	
2	1:05 PM	Review of September minutes	Alissa Craft	X
3	1:10 PM	Review Public Comment on draft coverage guidance on Osteoporosis Screening And Monitoring By Dual-Energy X-Ray Absorptiometry (DXA)	Alison Little	X
4	1:40 PM	Review Public Comment on draft coverage guidance on Treatment of Sleep Apnea for Adults	Alison Little	X
5	2:10 PM	Review initial draft coverage guidance on Intensity-Modulated Radiation Therapy	Wally Shaffer	X
6	3:10 PM	Review initial draft coverage guidance on Stereotactic Radiosurgery and Stereotactic Body Radiation Therapy	Wally Shaffer	X
7	3:50 PM	Confirm next meeting:	Alissa Craft	
8	3:55 PM	General Public Comment		
9	4:00 PM	Adjournment	Alissa Craft	

MINUTES

Health Technology Assessment Subcommittee
Meridian Park Community Health Education Center
19300 SW 65th Avenue, Tualatin, OR
September 23, 2013
1:00-4:00pm

Members Present: Alissa Craft, DO, MBA; Gerald Ahmann, MD; George Waldmann, MD; Timothy Keenen, MD (arrived 1:58 p.m.); Tracy Muday, MD.

Members Absent: James MacKay, MD

Staff Present: Darren Coffman; Wally Shaffer, MD, MPH; Jason Gingerich.

Also Attending: Alison Little, MD (CEBP); Shannon Vandergriff (CEBP); Denise Taray (DMAP); Andrea Herzka, , MD (OHSU); Nan Heim (Oregon Association of Orthopedists).

---

1. CALL TO ORDER

Alissa Craft called the meeting of the Health Technology Assessment Subcommittee (HTAS) to order at 1:00 pm.

---

2. MINUTES REVIEW

The subcomitteee requested the correction of a spelling error in the June minutes.
Minutes approved as amended 4-0 (Keenen not present).

---

3. Treatment of Sleep Apnea in Adults

Shaffer introduced the topic. This coverage guidance was recommended to HERC previously by HTAS before the GRADE methodology and HERC Coverage Guidance Development Framework were adopted. When the HERC reviewed the draft coverage guidance, they expressed concern about the weakness of evidence for coverage for surgical interventions. As that was being researched, staff found additional evidence regarding the levels of apnea hypopnea index at which CPAP offers clinically meaningful benefit. Shaffer said he has laid out three options for CPAP coverage in the meeting materials, with the first being the previously adopted language. In addition, the recommendation for coverage of surgery is suggested to be changed to a recommendation for noncoverage.

For surgeries the subcommittee discussed the high quality studies showing no benefit over sham surgery for radiofrequency ablation. Muday said she does not see many requests for radiofrequency ablation. Craft asked what surgeries the plans were seeing requests for. The subcommittee discussed which surgeries are frequently requested under the Oregon Health Plan. Waldmann said that tonsillectomy and adenoidectomy is common in children and not uncommon in adults. Muday said she sees more nasal surgeries. Sometimes they see requests

DRAFT

for combination surgery. She said that in some cases the surgeons say surgery won't treat the sleep apnea but will allow better tolerance of CPAP, but this question isn't addressed by the evidence.

Craft reviewed the HERC Coverage Development Framework as shown in the meeting materials leading to a "do not recommend" decision with a weak recommendation. The group voted 4-0 (Keenen not present) to accept the change in recommendation so that surgeries will not be recommended for coverage.

For CPAP, Waldmann asked Alison Little what the clinical outcomes were that the studies looked at. Little said clinical outcomes were considered to be effects on anything other than symptoms and gave as examples effects on diabetes and hypertension. Craft noted that use is quite prevalent so the subcommittee will need to weigh the evidence against the common use of these devices in practice. Craft reviewed the HERC Coverage Guidance Development Framework for the different AHI levels. For an AHI of 5-14 the framework would lead to a recommendation of noncoverage. For patients with an AHI of at least 15, the framework would recommend coverage. This would lead to option 2 from the meeting materials. Waldmann noted that the coverage would be more restrictive than Medicare. Shaffer noted that this may create implementation issues, but OHP deals with that at times. Craft pointed out that the coverage guidance would not apply to children.

The group decided to accept option 2 from the meeting materials, with minor punctuation changes to clarify under which conditions CPAP was recommended for coverage. Craft requested public comment but no one chose to comment. The motion was adopted 4-0 (Keenen not present), requesting that the draft coverage guidance with language as shown below be posted for public comment.

DRAFT

HERC COVERAGE GUIDANCE

Coverage of treatment for Obstructive Sleep Apnea (OSA) in adults should be limited, as follows:

CPAP should be 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 30 events per hour; or if between 15 and 30 events with additional symptoms including one or more of the following:
    - excessive daytime sleepiness (Epworth Sleepiness Scale score > 10), or
    - documented hypertension, or
    - ischemic heart disease, or
    - 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 should be 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.

Coverage of mandibular advancement devices (oral appliances) should be provided.

Intensive weight loss programs (if provided in the benefit package) should be covered for patients with obesity and obstructive sleep apnea.

Surgery for sleep apnea for adults should not be covered.

---

4. REVIEW OF PUBLIC COMMENTS

A) Upper Endoscopy for Gastroesophageal Reflux Disease (GERD) and Dyspepsia Symptoms

Little reviewed the public comments and the staff recommended responses in the meeting materials. Wally Shaffer said that in addition to the comments in the public comment disposition, he received feedback from CCO medical directors suggesting adding a requirement for counseling on lifestyle modifications. Since the subcommittee reviewed the previous draft, he changed the coverage guidance recommendation language to require advice on lifestyle modifications and a requirement for a trial of twice-daily PPI therapy. He said that the indication of troublesome dysphagia as an indication for endoscopy came from a high quality guideline and that the subcommittee may wish to add other indications or alarm symptoms.

Waldmann suggested it would be foolish not to do the endoscopy for anemia, weight loss or dysphagia because these could indicate gastric cancer. Craft suggested the "including but not limited to" language might put plan medical directors in a difficult situation. Muday said that if someone has acute anemia, it's fine. She said the guidance shouldn't be too prescriptive because endoscopy might be appropriate in cases such as a bleeding ulcer, which is treatable. Ahmann suggested limiting the anemia indication to "iron-deficient" anemia. Subcommittee members also requested that staff specify in all places that this guidance refers to upper endoscopy to avoid confusion with lower endoscopy, and decided to clarify that the recommendation for noncoverage of repeat endoscopy allow for an exception for patients whose previous upper endoscopy indicated Barrett's esophagus, since it is a pre-malignant condition.

A motion was made to approve the draft coverage guidance box language as modified below and to refer it to VbBS and HERC. **Motion approved 5-0.**

HERC COVERAGE GUIDANCE

Upper endoscopy for uninvestigated dyspepsia or GERD symptoms is not recommended for coverage in patients less than 50 years of age unless the patient has persistent symptoms following advice on lifestyle modifications and completion of an appropriate course of twice daily PPI therapy or an H. pylori test and treat protocol (*strong recommendation*).

Upper endoscopy for uninvestigated dyspepsia or GERD symptoms is recommended for coverage in patients at least 50 years of age (strong recommendation).

Upper endoscopy is recommended for coverage in patients with troublesome dysphagia, regardless of age or prior upper endoscopy.

DRAFT

In the absence of significant new symptoms, repeat upper endoscopy is not recommended for coverage for patients with dyspepsia or GERD after non-malignant findings (with the exception of Barrett's esophagus) on initial upper endoscopy (weak recommendation).

Note: This guidance does not apply to coverage of upper endoscopy for patients presenting with "alarm symptoms" including, but not limited to, iron deficiency anemia or weight loss.

B) Hip Surgery Procedures for Femoroacetabular Impingement Syndrome

Shaffer said that staff received a high volume of public comments and references supporting coverage and raising points relevant to the discussion. Staff has identified the key questions raised in these comments for subcommittee discussion. Shaffer reviewed these questions with the subcommittee and then Little read through the public comments and the CEbP's recommended responses, with discussion held until afterwards. Coffman invited Dr. Andrea Herzka, the appointed ad hoc expert for this topic, to come forward so she could address subcommittee members' questions.

Waldmann asked Herzka how long this surgery has been done. Herzka said that the condition was first described in 2003 but that the surgeon who described it had done some surgeries before that. Modern techniques became prevalent in 2005, increasing each year since then. Prior to that hip arthroscopy in general was mostly debridements as the techniques for labral repair had not been developed and the importance of labral tears was not known.

Waldmann asked about the training required for the procedure. Herzka said that the majority of hip arthroscopists are trained in a fellowship, though there is a physician at OHSU who performs the surgery and has not been trained at a fellowship. Waldmann expressed concern that an orthopedist without adequate training might start doing the procedure. Herzka said that you have to trust the physician to know the boundaries of their practice. There are hip arthroscopy replacement surgeons who won't do complex revisions. Keenen, an orthopedist, said that he has queried several orthopedists and they all refer patients with FAI to experts on this surgery because they don't want to go through the learning curve. He believes the likelihood of an "epidemic" of this surgery is low. Muday asked how many surgeons are doing this procedure in Oregon. Herzka said there are two surgeons including herself who do high volumes of this procedure and a couple of others who are becoming increasingly skilled.

Craft asked if anyone wanted to provide public comment. There was none.

Keenen said that he went through the algorithm before the meeting and got to the same place as Herzka—there is no acceptable alternative for those who have failed PT since these patients have already had nontreatment. He gave the example of herniated disks, which was identified in the 1920s. Before this, patients with herniated disks were previously identified due to symptoms such as foot drop rather than herniated disk. He admits that the impact of the surgery on progression to osteoarthritis is unknown, but believes the surgery has enough efficacy to warrant coverage—but we will never know if it prevents arthritis unless we cover it. In addition it would be difficult to recruit FAI patients for an RCT as the surgery is prevalent.

DRAFT

Subcommittee members discussed what other services payers might need to stop covering in order to pay for the cost of this service if it were added above the funding line on the Prioritized List for the Oregon Health Plan. Keenen argued that the successful treatment for FAI might displace costs for NSAIDs and the side effects of ulcers as well as the costs of people missing work and being unable to function in society. Therefore, it actually might result in long-term cost savings.

Muday brought up the case of hormone replacement therapy for menopause where treatment became prevalent based on lower quality evidence but was found to have more harms than benefits after additional research. Herzka responded that there haven't been significant adverse effects and the 8-year results look very stable, though we still don't know whether the surgery delays osteoarthritis. She believes the worst case is medium term pain relief and no delay in arthritis onset and the best case is it would prevent arthritis.

Craft asked the subcommittee whether there was a motion based on the public comment heard, to which there was a motion and a second to change the guidance to a recommendation for coverage, at which point discussion followed. The subcommittee discussed how covering the procedure with restrictions, including age 15-55 and a failure of physical therapy, might be appropriate. Staff then presented some coverage limitations based on Herzka's recommendations and those of some commercial payers which cover the procedure.

The group reviewed the coverage language and made a few changes, changing the age limitation to up to age 55, and requiring failure of physical therapy and additional noninvasive treatments before coverage would be authorized. Craft pointed out that for this condition, OHP may not cover physical therapy.

After discussion the subcommittee agreed to change the algorithm to show insufficient/mixed evidence without alternative treatments available (after a failure of conservative management), with treatment risk similar or less than alternatives and higher quality trials not reasonable, leading to a strong recommendation for coverage.

Little asked why a sham surgery isn't an option. Herzka said that orthopedists wouldn't feel it was ethical to make an incision and not repair the damage, resulting in a patient who might get permanent disability in their thirties due to inability to sit, drive or stand. Waldmann said he believes it could be done but would not likely be done. Keenen cited a similar situation where a surgery was attempted but there was a high rate of crossover between the treatment and control, making the results of limited use.

Gingerich pointed out that the public comment disposition would require significant revisions and the subcommittee requested that staff make such revisions based on subcommittee discussion. The HERC Coverage Guidance Development Framework will also be modified to reflect the lack of available alternatives, treatment risk compared to no treatment is similar or less, treatment is prevalent and a research study is not reasonable, with a strong recommendation for coverage.

The subcommittee voted 4-0 (Muday abstained) to refer the coverage guidance with the language below to VbBS and HERC.

DRAFT

HERC COVERAGE GUIDANCE

Surgery for femoroacetabular impingement (FAI) syndrome is recommended *(strong recommendation)* for coverage in patients who meet all of the following criteria:

- Adult patients who are younger than 55 years of age, or adolescent patients who are skeletally mature with documented closure of growth plates; and
- Other sources of pain have been ruled out (e.g., lumbar spine pathology, SI joint dysfunction, sports hernia); and
- Pain unresponsive to physical therapy and other non-surgical management and conservative treatments (e.g., restricted activity, cortisone injections, nonsteroidal anti-inflammatory drugs) of at least three months duration, or conservative therapy is contraindicated; and
- Moderate-to-severe persistent hip or groin pain that significantly limits activity and is worsened by flexion activities (e.g., squatting or prolonged sitting); and
- Positive impingement sign (i.e., sudden pain on 90 degree hip flexion with adduction and internal rotation or extension and external rotation); and
- Radiographic confirmation of FAI morphology (e.g., pistol-grip deformity, alpha angle greater than 50 degrees, coxa profunda, and/or acetabular retroversion); and
- Do not have advanced osteoarthritis (i.e., Tönnis* grade 2 or 3) or severe cartilage damage (i.e., Outerbridge* grade III or IV).

*See Appendix D for description of Tönnis and Outerbridge grading systems.

Appendix D. Grading Systems

Tonnis Classification of Osteoarthritis by Radiographic Changes

- Grade 0: No signs of osteoarthritis
- Grade 1: Increased sclerosis of femoral head or acetabulum, slight joint space narrowing or slight slipping of joint margin, no or slight loss of head sphericity
- Grade 2: Small cysts in femoral head or acetabulum, moderate joint space narrowing, moderate loss of head sphericity
- Grade 3: Large cysts, severe joint space narrowing or obliteration of joint space, severe deformity of the head, avascular necrosis

Outerbridge grades include the following:

- Grade 0: Normal
- Grade I: Cartilage with softening and swelling
- Grade II: Partial-thickness defect with fissures on the surface that do not reach subchondral bone or exceed 1.5 cm in diameter
- Grade III: Fissuring to the level of subchondral bone in an area with a diameter more than 1.5 cm
- Grade IV: Exposed subchondral bone head

C) Osteoporosis Screening And Monitoring By Dual-Energy X-Ray Absorptiometry (DXA)

DRAFT

The subcommittee deferred discussion of this topic to its next meeting due to time constraints.

---

5. NEW TOPICS

The subcommittee discussed several options for new coverage guidance topics for its next meeting, and agreed that prophylactic mastectomy, oncotype diagnostic testing for cancer and insulin pumps for diabetes would all be appropriate topics for future discussion. Waldmann suggested staff may wish to wait for the results of the TAILORx trial on oncotype testing before taking on that topic.

---

6. PUBLIC COMMENT

Craft offered an additional opportunity for members of the public to comment but no one chose to comment.

---

7. ADJOURNMENT

The meeting was adjourned at 4:00 pm. The next meeting is scheduled for November 25 from 1:00-4:00pm in Room 117B of the Meridian Park Hospital Community Health Education Center in Tualatin.

DRAFT

Section 2.0

Coverage Guidances

HEALTH EVIDENCE REVIEW COMMISSION (HERC)

COVERAGE GUIDANCE: OSTEOPOROSIS SCREENING AND MONITORING BY DUAL-ENERGY X-RAY ABSORPTIOMETRY (DXA)

DRAFT for HTAS Meeting Materials 11/25/2013

HERC COVERAGE GUIDANCE

Osteoporosis screening by dual-energy X-ray absorptiometry (DXA) is recommended for coverage only for women aged 65 or older, and for men or younger women whose fracture risk is equal to or greater than that of a 65 year old white woman who has no additional risk factors. Fracture risk should be assessed by the World Health Organization's FRAX tool or similar instrument (*strong recommendation*).

Repeat osteoporosis screening by DXA, for women with normal bone density, is not recommended for coverage more frequently than once every fifteen years (*weak recommendation*).

Routine osteoporosis screening by DXA is not recommended for coverage in men (*weak recommendation*).

~~For~~ Unless there has been significant change in the individual's risk factors, such that rapid changes in bone density are expected, monitoring of individuals with low bone ~~mineral~~ density~~, monitoring~~ by repeat DXA scanning is ~~not~~ recommended for coverage ~~more often than~~ (weak recommendation) only at the following frequencies:

- once every two years for those with osteoporosis or advanced osteopenia (T-score of -2.00 or lower~~),~~)
- once every four years for moderate osteopenia (T-score between -1.50 and -1.99~~), and~~)
- once every fifteen years for mild osteopenia (T-score between -1.01 and -1.49~~), unless there has been significant change in the individual's risk factors.~~).

Repeat testing should only be covered if the results will influence clinical management~~ or if rapid changes in bone density are expected (*weak recommendation*).~~. For purposes of monitoring osteoporosis medication therapy, testing at intervals of less than two years is not recommended for coverage (*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

DRAFT

- 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 Heath 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

Gourlay, M.L., Fine, J.P., Preisser, J.S., May, R.C., Li, C., Lui, L., et al. (2012). Bone-density testing interval and transition to osteoporosis in older women. *New England Journal of Medicine, 366*(3), 225-233.

National Clinical Guideline Center. (2012). *Osteoporosis: Assessing the risk of fragility fracture*. London: National Clinical Guideline Center. Retrieved May 10, 2013, from http://guidance.nice.org.uk/CG146/Guidance

Nelson, H.D., Haney, E.M., Chou, R., Dana, T., Fu, R., & Bougatsos, C. (2010). *Screening for osteoporosis: Systematic review to update the 2002 U.S. Preventive Services Task Force recommendation*. Evidence Synthesis No. 77. AHRQ Publication No. 10-05145-EF-1. Rockville, Maryland: Agency for Healthcare Research and Quality. Retrieved May 10, 2013, from http://www.ncbi.nlm.nih.gov/books/NBK45201/

U.S. Preventive Services Task Force. (2011). Screening for osteoporosis: U.S. Preventive Services Task Force recommendation statement. *Annals of Internal Medicine, 154*(5), 356-364. Retrieved May 10, 2013, from http://www.uspreventiveservicestaskforce.org/uspstf/uspsoste.htm

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

SUMMARY OF EVIDENCE

Clinical Background

Osteoporosis is characterized by low bone mineral density (BMD) and a resultant increased risk for fractures. It is estimated that as many as 1 in 2 women and 1 in 5 men are at risk for an osteoporosis-related fracture during their lifetime. Osteoporosis is more common in women than men and is more common in white persons than in any other racial group. For all demographic groups, the rates of osteoporosis increase with age. Elderly patients have increased susceptibility to fractures because they commonly have

additional risk factors for fractures, such as poor bone quality and an increased tendency to fall. Hip fractures in particular can result in significant morbidity and mortality. Fractures at other sites also can lead to significant illness, causing chronic pain or disability and negatively affecting functional ability and quality of life. Direct medical care costs of osteoporotic fractures were estimated to be $12.2 to $17.9 billion per year in 2002 U.S. dollars; these estimates do not include indirect costs associated with lost productivity of patients and caregivers.

Many different risk assessment instruments have been developed to predict risk for low BMD or fractures. Multiple studies have validated these tools; however, few of these studies have included men. Despite various risk factors and variables included in the different risk assessment tools, none of the tools has consistently superior performance. The FRAX tool, developed by the World Health Organization and the National Osteoporosis Foundation, is one of the most widely used instruments to predict risk for fractures. This tool was derived from data on 9 cohorts in Europe, Canada, the United States, and Japan. Seven of these cohorts included men. The FRAX tool was validated in 11 cohorts, but only 1 of these cohorts included men. Because a large and diverse sample was used to develop and validate the FRAX tool and this instrument includes a publicly available risk calculator, the USPSTF used the FRAX tool to determine which individuals would exceed the baseline risk threshold for fractures on the basis of their age or other risk factors (such as low BMI, parental history of hip fracture, smoking status, and daily alcohol use). Considering a 65-year-old white woman who has no other risk factors to be the baseline risk case (a 10-year risk for any osteoporotic fracture of 9.3%), women as young as 50 years may have a 10-year risk for any osteoporotic fracture of 9.3% or greater, depending on the type and number of risk factors present.

Bone mineral density (BMD) criteria were developed by the World Health Organization (WHO) from epidemiologic data that describe the normal distribution of BMD in a young healthy reference population. Osteoporosis is diagnosed when the BMD at the spine, hip, or wrist is 2.5 or more standard deviations (SD) below the reference mean. Low bone density or mass (sometimes referred to as osteopenia) is diagnosed when BMD is between 1.0–2.5 SD below the reference mean. The number of standard deviation units above or below the young healthy mean is called the T-score. Although intended for epidemiologic purposes, T-scores have been used as selection criteria for trials of therapies. They are now used to identify individuals with low BMD and to make treatment decisions.

Evidence Review

USPSTF

Detection

The USPSTF found convincing evidence that bone measurement tests predict short-term risk for osteoporotic fractures in women and men. The most commonly used tests are dual-energy x-ray absorptiometry (DXA) of the hip and lumbar spine and quantitative ultrasonography of the calcaneus. Adequate evidence indicates that clinical risk assessment instruments have only modest predictive value for low bone density or fractures.

Benefits of Detection and Early Intervention

No controlled studies have evaluated the effect of screening for osteoporosis on fracture rates or fracture related morbidity or mortality. In postmenopausal women who have no previous osteoporotic fractures, the USPSTF found convincing evidence that drug therapies reduce the risk for fractures. In women aged 65 years or older and in younger women whose fracture risk is equal to or greater than that of a 65-year-old white woman who has no additional risk factors, the USPSTF judged that the benefit of treating screening-detected osteoporosis is at least moderate. Because of the lack of relevant studies, the USPSTF found inadequate evidence that drug therapies reduce the risk for fractures in men who have no previous osteoporotic fractures.

Accuracy of Screening Tests

DXA

Measurement of bone density using DXA has become the gold standard for the diagnosis of osteoporosis and for guiding decisions about which patients to treat. Although it is not a perfect predictor of fractures, DXA of the femoral neck is considered the best predictor of hip fracture and is comparable with DXA measurements of the forearm for predicting fractures at other sites. Previous studies evaluating the accuracy of DXA for predicting fractures have focused mainly on women; studies have only recently assessed the predictive ability of DXA in men. A large prospective cohort study in the Netherlands that included men and women older than 55 years reported the incidence of vertebral and nonvertebral fractures approximately 6 years after baseline DXA measurements of the femoral neck were obtained. For each SD reduction in BMD at the femoral neck, the hazard ratio for vertebral and non-vertebral fractures increased to a similar degree in both men and women. Other studies of the performance of DXA in men have reported similar findings.

Quantitative Ultrasonography

The most commonly used test in the United States after DXA is quantitative ultrasonography (US) of the calcaneus. Quantitative US is less expensive than DXA,

DRAFT

does not involve radiation, and can feasibly be implemented in primary care settings. Recent studies demonstrate that quantitative US of the calcaneus can predict fractures as effectively as DXA in postmenopausal women and in men. Quantitative US seems to be equivalent to DXA for predicting fractures and has other potential advantages, but also a few distinct disadvantages. The current diagnostic criteria for osteoporosis use DXA measurements as cutoffs, and the measurements obtained from quantitative US are not interchangeable with those obtained from DXA. Also, all trials evaluating drug therapies for osteoporosis use DXA measurements as inclusion criteria. Thus, for quantitative US to be relevant and clinically useful, a method for converting or adapting results of quantitative US to the DXA scale will need to be developed.

One meta-analysis examined 25 studies to assess the accuracy of quantitative US compared with DXA in identifying patients with osteoporosis. When various quantitative US index parameter cutoffs were used, the results varied widely in sensitivity and specificity for identifying individuals with a T-score of -2.5 or less on DXA. No quantitative US cutoff existed at which sensitivity and specificity were both high.

Frequency of Monitoring

The USPSTF did not make any specific recommendations regarding screening interval or frequency. The systematic review conducted to support the recommendation reported on only one study that addressed this question, a large good-quality prospective cohort study of 4,124 women age ≥65 years from the Study of Osteoporotic Fractures. This study found that repeating a BMD measurement up to 8 years after an initial measurement did not significantly change estimates for non-vertebral, hip, or vertebral fractures. No studies of screening intervals have been conducted in men or other groups of women.

Because of the limited evidence supporting frequency of monitoring, an additional search of the literature was undertaken from the end date of the Nelson review (December 2009). One study was identified that addressed frequency of monitoring (Gourlay et al. 2012). This NIH funded study evaluated women with normal or osteopenic BMD who were older than 66 years of age and had no history of hip or vertebral fracture. Osteopenia was categorized as mild (T-score -1.01 to -1.49), moderate (T-score -1.50 to -1.99) or advanced (T-score -2.0 to -2.49). They were followed prospectively for 15 years and the BMD testing interval, defined as the estimated time for 10% of women to make the transition to osteoporosis, was calculated. The estimated BMD testing interval was 16.8 years (95% CI, 11.5 to 24.6) for women with normal BMD, 17.3 years (95% CI, 13.9 to 21.5) for women with mild osteopenia, 4.7 years (95% CI, 4.2 to 5.2) for women with moderate osteopenia, and 1.1 years (95% CI, 1.0 to 1.3) for women with advanced osteopenia.

Effectiveness of Early Detection and Treatment

No controlled studies have evaluated the effect of screening for osteoporosis on rates of fractures or fracture related morbidity or mortality. Drug therapies for osteoporosis can be for primary prevention (prevention of an osteoporotic fracture in patients with low BMD who have no previous fractures) or secondary prevention (prevention of an osteoporotic fracture in patients who have a known previous osteoporotic fracture). Primary prevention trials are more applicable to the screening population addressed in this recommendation. Drug therapies include bisphosphonates, parathyroid hormone, raloxifene, estrogen, and calcitonin. For primary prevention in postmenopausal women, bisphosphonates, parathyroid hormone, raloxifene, and estrogen have been shown to reduce vertebral fractures. The evidence is strongest and most consistent for bisphosphonates and raloxifene.

In a meta-analysis of 7 trials, the relative risk (RR) for vertebral fractures for bisphosphonates compared with placebo was 0.66 (95% CI, 0.50 to 0.89). Two large placebo controlled trials of raloxifene reported reduced vertebral fractures, with a combined RR for raloxifene of 0.61 compared with placebo (CI, 0.55 to 0.69). A pooled analysis of 9 trials demonstrated a non–statistically significant trend toward a reduction in non-vertebral fractures with bisphosphonates compared with placebo (RR, 0.83 [CI, 0.64 to 1.08]). In the largest trial of bisphosphonates, the Fracture Intervention Trial of alendronate, fractures were significantly reduced only in women with baseline femoral neck T-scores less than -2.5. Evidence of the effectiveness of treatment of osteoporosis in men is limited. There are no primary prevention trials of bisphosphonates in men and only 2 secondary prevention trials of alendronate. When the 2 trials were pooled, alendronate was associated with a reduced risk for vertebral fractures (odds ratio [OR], 0.35 [CI, 0.17 to 0.77]), and the effect on non-vertebral fractures was not statistically significant (OR, 0.73 [CI, 0.32 to 1.67]). A single primary prevention trial of parathyroid hormone in men reported a non-statistically significant trend toward a reduction in vertebral and non-vertebral fractures. None of the other therapies for osteoporosis in men has been evaluated in randomized trials.

Potential Harms of Screening and Treatment

Potential harms of screening for osteoporosis include false-positive test results causing unnecessary treatment, false-negative test results, and patient anxiety about positive test results. No studies that addressed the potential harms of screening were identified during this review. The harms of drug therapy for osteoporosis have been studied most extensively for bisphosphonates, raloxifene, and estrogen. For bisphosphonates, the evidence demonstrates no definitive increase in the risk for serious gastrointestinal adverse events (for example, perforations, ulcers, bleeding, esophagitis, or esophageal ulceration) in persons who use these medications appropriately. The evidence on the risk for atrial fibrillation with bisphosphonates is conflicting. One large case-control study

in Denmark showed an increased risk for atrial fibrillation with any use of alendronate compared with no use of this agent (OR, 1.86 [CI, 1.09 to 3.15]), but a smaller case–control study in Washington showed no increased risk for atrial fibrillation with any use of etidronate (RR, 0.95 [CI, 0.84 to 1.07]) or any use of alendronate (RR, 1.04 [CI, 0.90 to 1.21]) compared with no use of either agent.

Osteonecrosis of the jaw has been associated with bisphosphonates in case reports, but this condition typically develops in patients with cancer who receive higher doses than those normally used for osteoporosis treatment or prevention. Case reports also have described severe musculoskeletal symptoms associated with all of the bisphosphonates. In October 2010, the U.S. Food and Drug Administration issued a warning about a possible elevated risk for midfemur fractures in patients receiving bisphosphonates, especially for patients who have received them for more than 5 years.

Raloxifene and estrogen are associated with higher rates of thromboembolic events than placebo. Estrogen increases the risk for stroke, and estrogen with progestin increases the risk for coronary heart disease and breast cancer. Evidence is limited on the harms associated with use of calcitonin and parathyroid hormone for osteoporosis.

Overall, the USPSTF found no new studies that described harms of screening for osteoporosis in men or women. Screening with DXA is associated with opportunity costs (time and effort required by patients and the health care system). Harms of drug therapies for osteoporosis depend on the specific medication used. The USPSTF found adequate evidence that the harms of bisphosphonates, the most commonly prescribed therapies, are no greater than small. Convincing evidence indicates that the harms of estrogen and selective estrogen receptor modulators are small to moderate.

Estimate of Magnitude of Net Benefit

The USPSTF found convincing evidence that drug therapies reduce subsequent fracture rates in postmenopausal women. For women aged 65 years or older and younger women who have similar estimates of fracture risk, the benefit of treating screening-detected osteoporosis is at least moderate. The harms of treatment were found to range from no greater than small for bisphosphonates and parathyroid hormone to small to moderate for raloxifene and estrogen. Therefore, the USPSTF concludes with moderate certainty that the net benefit of screening for osteoporosis in this group of women is at least moderate. For men, the USPSTF concludes that evidence is inadequate to assess the effectiveness of drug therapies in reducing subsequent fracture rates in men who have no previous fractures. Treatments that have been proven effective in women cannot necessarily be presumed to have similar effectiveness in men. Thus, the USPSTF could not assess the balance of benefits and harms of screening for osteoporosis in men.

Overall USPSTF Assessment
The USPSTF concludes that for women aged 65 years or older and younger women whose fracture risk is equal to or greater than that of a 65-year-old white woman who has no additional risk factors, there is moderate certainty that the net benefit of screening for osteoporosis by using DXA is at least moderate. The USPSTF concludes that for men, evidence of the benefits of screening for osteoporosis is lacking and the balance of benefits and harms cannot be determined.

[Evidence Source]

NICE GUIDELINE

The NICE guideline makes the follow recommendations pertaining to assessing the risk of fragility fractures:

Targeting risk assessment
1. Consider assessment of fracture risk:

- in all women aged 65 years and over and all men aged 75 years and over
- in women aged under 65 years and men aged under 75 years in the presence of risk factors, for example:
  - previous fragility fracture,
  - current use or frequent recent use of oral or systemic glucocorticoids,
  - history of falls,
  - family history of hip fracture,
  - other causes of secondary osteoporosis[1],
  - low body mass index (BMI) (less than 18.5 kg/m2),
  - smoking,
  - alcohol intake of more than 14 units per week for women and more than 21 units per week for men.

2. Do not routinely assess fracture risk in people aged under 50 years unless they have major risk factors (for example, current or frequent recent use of oral or systemic glucocorticoids, untreated premature menopause or previous fragility fracture), because they are unlikely to be at high risk.

[1] Causes of secondary osteoporosis include endocrine (hypogonadism in either sex including untreated premature menopause and treatment with aromatase inhibitors or androgen deprivation therapy; hyperthyroidism; hyperparathyroidism; hyperprolactinaemia; Cushing's disease; diabetes), gastrointestinal (coeliac disease; inflammatory bowel disease; chronic liver disease; chronic pancreatitis; other causes of malabsorption), rheumatological (rheumatoid arthritis; other inflammatory arthropathies), haematological (multiple myeloma; haemoglobinopathies; systemic mastocytosis), respiratory (cystic fibrosis; chronic obstructive pulmonary disease), metabolic (homocystinuria), chronic renal disease and immobility(due for example to neurological injury or disease).

3. Estimate absolute risk when assessing risk of fracture (for example, the predicted risk of major osteoporotic or hip fracture over 10 years, expressed as a percentage).

4. Use either FRAX[2] (without a bone mineral density [BMD] value, if a dual-energy X-ray absorptiometry [DXA] scan has not previously been undertaken) or QFracture[3], within their allowed age ranges, to estimate 10-year predicted absolute fracture risk when assessing risk of fracture. Above the upper age limits defined by the tools, consider people to be at high risk.

5. Interpret the estimated absolute risk of fracture in people aged over 80 years with caution, because predicted 10-year fracture risk may underestimate their short-term fracture risk.

6. Do not routinely measure BMD to assess fracture risk without prior assessment using FRAX (without a BMD value) or QFracture.

7. Following risk assessment with FRAX (without a BMD value) or QFracture, consider measuring BMD with DXA in people whose fracture risk is in the region of an intervention threshold[4] for a proposed treatment, and recalculate absolute risk using FRAX with the BMD value.

8. Consider measuring BMD with DXA before starting treatments that may have a rapid adverse effect on bone density (for example, sex hormone deprivation for treatment for breast or prostate cancer).

9. Measure BMD to assess fracture risk in people aged under 40 years who have a major risk factor, such as history of multiple fragility fracture, major osteoporotic fracture, or current or recent use of high-dose oral or systemic glucocorticoids (more than 7.5 mg prednisolone or equivalent per day for 3 months or longer).

10. Consider recalculating fracture risk in the future:

[2] FRAX, the WHO fracture risk assessment tool, is available from www.shef.ac.uk/FRAX. It can be used for people aged between 40 and 90 years, either with or without BMD values, as specified.

[3] QFracture is available from www.qfracture.org. It can be used for people aged between 30 and 84 years. BMD values cannot be incorporated into the risk algorithm.

[4] An intervention threshold is the level of risk at which an intervention is recommended. People whose risk is in the region from just below to just above the threshold may be reclassified if BMD is added to assessment. It is out of the scope of this guideline to recommend intervention thresholds. Healthcare professionals should follow local protocols or other national guidelines for advice on intervention thresholds.

- if the original calculated risk was in the region of the intervention threshold[5] for a proposed treatment and only after a minimum of 2 years, or
- when there has been a change in the person's risk factors.

11. Take into account that risk assessment tools may underestimate fracture risk in certain circumstances, for example if a person:

- has a history of multiple fractures
- has had previous vertebral fracture(s)
- has a high alcohol intake
- is taking high-dose oral or high-dose systemic glucocorticoids (more than 7.5 mg prednisolone or equivalent per day for 3 months or longer)
- has other causes of secondary osteoporosis.[6]

12. Take into account that fracture risk can be affected by factors that may not be included in the risk tool, for example living in a care home or taking drugs that may impair bone metabolism (such as anti-convulsants, selective serotonin reuptake inhibitors, thiazolidinediones, proton pump inhibitors and anti-retroviral drugs).

[Evidence Source]

Evidence Summary

Bone measurement tests predict short-term risk for osteoporotic fractures in women and men. The most appropriate interval for screening has not been identified, but repeating a BMD measurement up to 8 years after an initial measurement does not significantly change fracture estimates, and transition to osteoporosis occurs for most women with normal BMD no sooner than 17 years. In postmenopausal women who have no previous osteoporotic fractures, drug therapies reduce the risk for fractures (primary prevention). Bisphosphonates, parathyroid hormone, raloxifene, and estrogen have all been shown to reduce vertebral fractures in this population. Potential harms of

[5] An intervention threshold is the level of risk at which an intervention is recommended. It is out of the scope of this guideline to recommend intervention thresholds. Healthcare professionals should follow local protocols or other national guidelines for advice on intervention thresholds.

[6] Causes of secondary osteoporosis include: endocrine (hypogonadism in either sex including untreated premature menopause and treatment with aromatase inhibitors or androgen deprivation therapy; hyperthyroidism; hyperparathyroidism; hyperprolactinaemia; Cushing's disease; diabetes), gastrointestinal (coeliac disease; inflammatory bowel disease; chronic liver disease; chronic pancreatitis; other causes of malabsorption), rheumatological (rheumatoid arthritis; other inflammatory arthropathies), haematological (multiple myeloma; haemoglobinopathies; systemic mastocytosis), respiratory (cystic fibrosis; chronic obstructive pulmonary disease), metabolic (homocystinuria), chronic renal disease and immobility (due for example to neurological injury or disease).

screening for osteoporosis include false-positive test results causing unnecessary treatment, false-negative test results, and patient anxiety about positive test results.

For women aged 65 years or older and younger women who have similar estimates of fracture risk, the benefit of treating screening-detected osteoporosis is at least moderate, while the harms range from small to moderate. Therefore, the net benefit of screening for osteoporosis in this group of women is at least moderate. For men, the evidence is inadequate to assess the effectiveness of drug therapies in reducing subsequent fracture rates in men who have no previous fractures.

DRAFT

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	Balance between desirable and undesirable effects	Quality of evidence*	Resource Allocation	Values and preferences	Coverage Recommendation
Screening for osteoporosis in women aged 65 or over, or with equivalent risks	Small to moderate net benefit	High	Moderately high on a population-wide basis, but with significant offsets if effective fracture prevention	Low variability (most people would prefer screening and fracture prevention)	Recommended for coverage (strong recommendation)
Screening for osteoporosis in men aged 70 or over	Unknown	Very low	Moderately high	Moderate variability (some would prefer availability of screening even if benefit not established)	Not recommended for coverage (weak recommendation)
Repeat DXA < 2 years for monitoring osteoporosis or advanced osteopenia	Likely no net benefit	Very low	Moderately significant cost associated with more frequent monitoring	Low variability	Not recommended for coverage (weak recommendation)

DRAFT

Indication	Balance between desirable and undesirable effects	Quality of evidence*	Resource Allocation	Values and preferences	Coverage Recommendation
Repeat DXA < 4 years for monitoring moderate osteopenia	Likely no net benefit	Very low	Moderately significant cost associated with more frequent monitoring	Low variability	Not recommended for coverage (weak recommendation)
Repeat screening DXA < 15 years in women with normal BMD or mild osteopenia	Likely no net benefit	Very low	Moderately significant cost associated with more frequent monitoring	Low variability	Not recommended for coverage (weak recommendation)

*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

DRAFT

POLICY LANDSCAPE

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, nine medical specialty organizations, along with Consumer Reports, have identified five tests or procedures commonly used in their field, whose necessity should be questioned and discussed. The American College of Rheumatology makes the following recommendation:

> **Don't routinely repeat DXA scans more often than once every two years.** Initial screening for osteoporosis should be performed according to National Osteoporosis Foundation recommendations. The optimal interval for repeating Dual-energy X-ray Absorptiometry (DXA) scans is uncertain, but because changes in bone density over short intervals are often smaller than the measurement error of most DXA scanners, frequent testing (e.g., <2 years) is unnecessary in most patients. Even in high-risk patients receiving drug therapy for osteoporosis, DXA changes do not always correlate with probability of fracture. Therefore, DXAs should only be repeated if the result will influence clinical management or if rapid changes in bone density are expected. Recent evidence also suggests that healthy women age 67 and older with normal bone mass may not need additional DXA testing for up to ten years provided osteoporosis risk factors do not significantly change.

Five quality measures were identified pertaining to BMD testing when searching the National Quality Measures Clearinghouse. All five were developed by the National Committee for Quality Assurance, and four of the five are endorsed by the NQF:

- Osteoporosis management in women who had a fracture: percentage of women 67 years of age and older who suffered a fracture and who had either a bone mineral density (BMD) test or prescription for a drug to treat or prevent osteoporosis in the six months after the fracture.
- Osteoporosis testing in older women: the percentage of Medicare women 65 years of age and over who report ever having received a bone density test to check for osteoporosis.
- Osteoporosis: percentage of patients aged 50 years and older with a fracture of the hip, spine or distal radius who had a central DXA measurement ordered or performed or pharmacologic therapy prescribed.
- Osteoporosis: percentage of female patients aged 65 years and older who have a central DXA measurement ordered or performed at least once since age 60 or pharmacologic therapy prescribed within 12 months.

The fifth measure has not been endorsed by the NQF:

- Osteoporosis: percentage of patients aged 18 years and older with one of the following conditions or therapies: receiving oral glucocorticosteroid therapy for greater than 3 months OR hypogonadism OR fracture history OR transplant history OR obesity surgery OR malabsorption disease OR receiving aromatase therapy for breast cancer who had a central dual-energy X-ray absorptiometry ordered or performed or pharmacologic therapy prescribed within 12 months.

COMMITTEE DELIBERATIONS – HTAS

COMMITTEE DELIBERATIONS – VBBS

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 of evidence across studies for the treatment/outcome

High = Further research is very unlikely to change our confidence in the estimate of effect.

Moderate = Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

Low = Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low = Any estimate of effect is very uncertain.

DRAFT

Appendix B. Applicable Codes

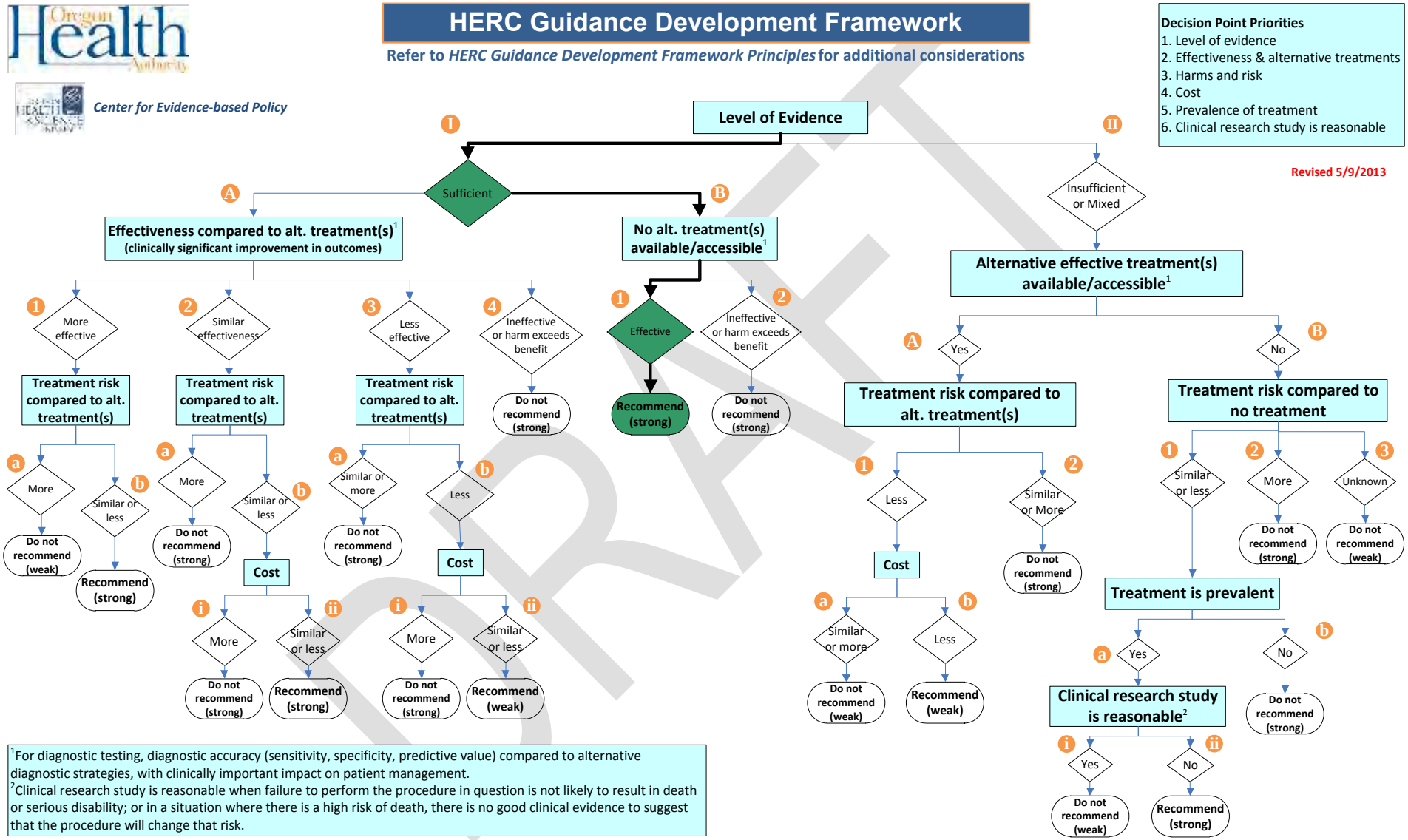
CODES	DESCRIPTION
ICD-9 Diagnosis Codes	
733.0	Osteoporosis
733.90	Disorder of bone and cartilage, unspecified
V82.81	Special screening for osteoporosis
ICD-9 Volume 3 (Procedure Codes)	
None	
CPT Codes	
76977	Ultrasound bone density measurement and interpretation, peripheral sites, any method
77080	Dual-energy X-ray absorptiometry (DXA), bone density study, 1 or more sites; axial skeleton (e.g., hips, pelvis, spine)
77081	Dual-energy X-ray absorptiometry (DXA), bone density study, 1 or more sites; appendicular skeleton (peripheral) (e.g., radius, wrist, heel)
HCPCS Level II Codes	
None	

Note: Inclusion on this list does not guarantee coverage

DRAFT

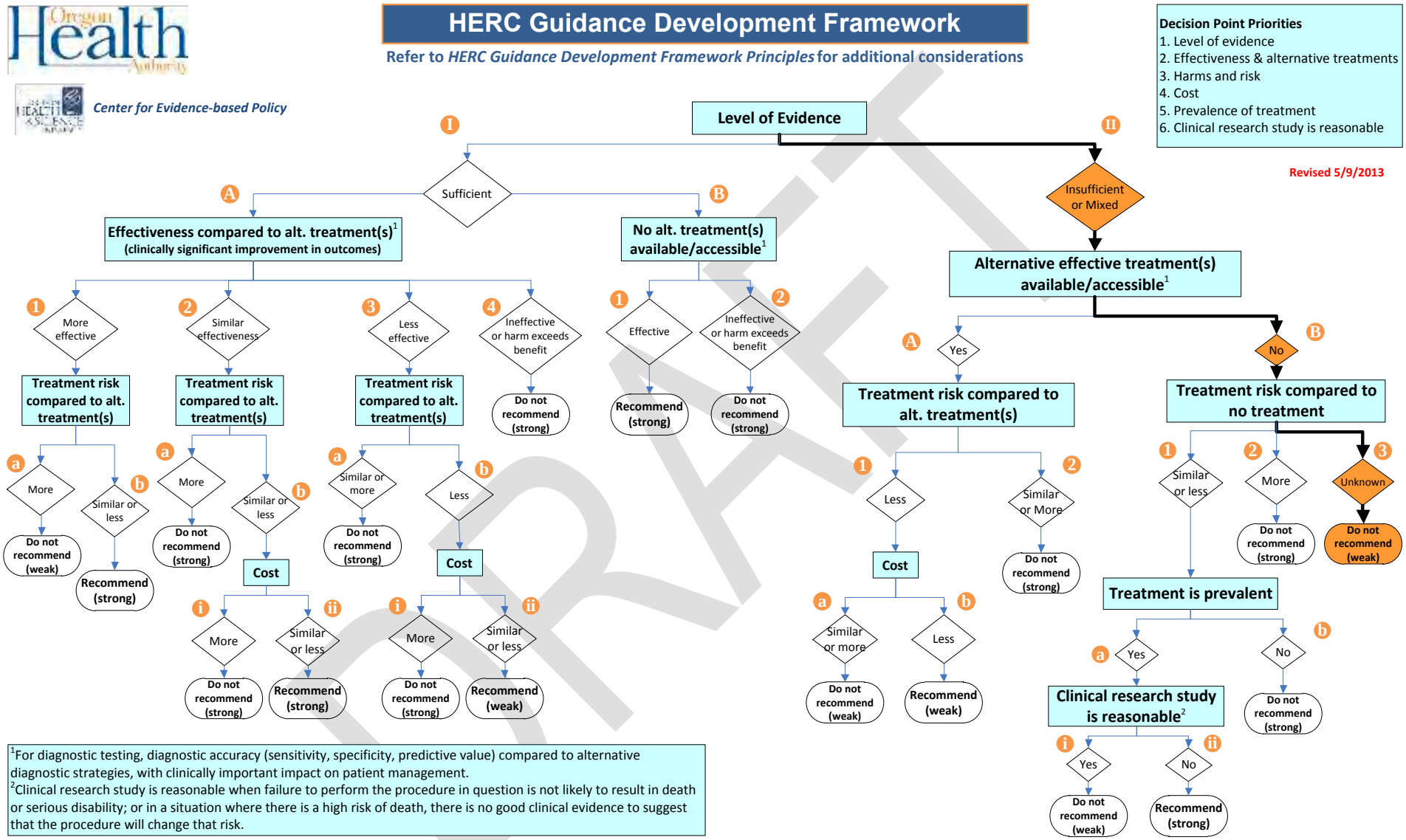
Appendix C. HERC Guidance Development Framework

Screening for osteoporosis in women aged 65 or over, or with equivalent risks



DRAFT

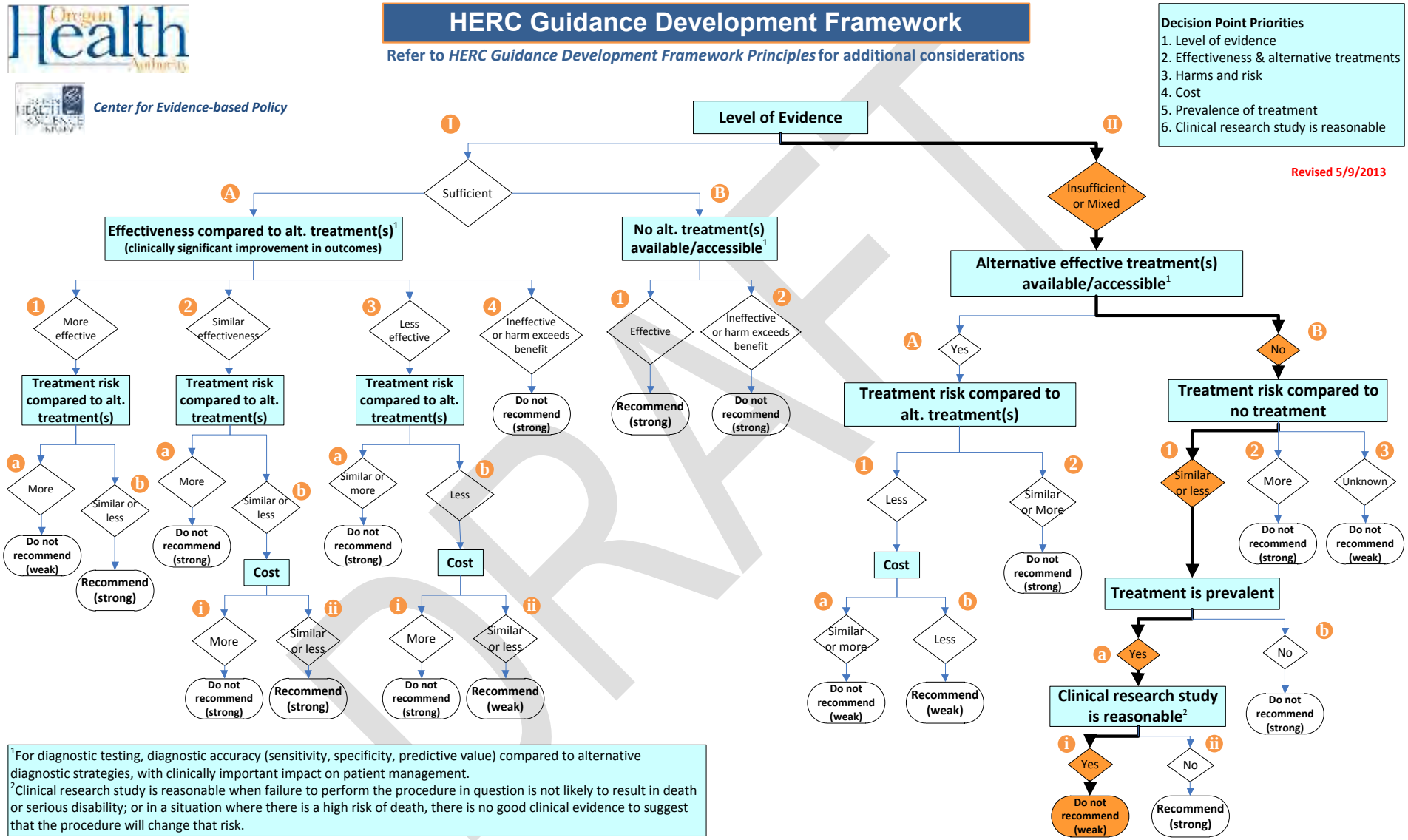
Screening for osteoporosis in men without additional risk factors



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

DRAFT

Repeat DXA for monitoring osteoporosis or advanced osteopenia < 2 years; Repeat screening <4 years for moderate, Repeat screening DXA < 15 years in women with normal BMD or mild osteopenia



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

DRAFT

HERC Coverage Guidance – Osteoporosis Screening And Monitoring By Dual-Energy X-Ray Absorptiometry (DXA)

Disposition of Public Comments

Expert	#	Comment	Disposition
E. Michael Lewiecki, MD, FACP, FACE Osteoporosis Director, New Mexico Clinical Research & Osteoporosis Center, Inc., Albuquerque, NM	1	**Background.** Osteoporosis, defined as low bone strength that increases the risk of fractures (1), is a common skeletal disorder that has been identified by the US Surgeon General as a major public health concern (2). About one of every two women and one of every five men will have an osteoporotic fracture in their lifetimes. Osteoporotic fractures are associated with an increase in morbidity and mortality, as well as high healthcare expenses (2). We are fortunately able to easily and inexpensively measure bone mineral density (BMD) by dual-energy X-ray absorptiometry (DXA) (3), assess fracture risk (4), and treat with pharmacological agents to reduce fracture risk (5). However, osteoporosis continues to be underdiagnosed (6) and undertreated (7), with those for whom treatment is started commonly failing to take medication correctly or long enough to achieve the expected benefit (8). This "treatment gap," the difference between the number of patients who could benefit from treatment and those who actually receive it (9), has created the need for better strategies to reduce the burden of osteoporotic fractures.	Thank you for this background information.
HERC-appointed Expert	2	**Clinical applications of DXA.** DXA is used to measure BMD, predict fracture risk, and monitor the skeletal effects of osteoporosis treatment (10). The National Osteoporosis Foundation (NOF) has developed evidence-based clinical practice guidelines, endorsed by numerous profession societies and updated in 2013, that provide clinicians with indications for BMD testing, treatment of osteoporosis, and monitoring treatment (11). The NOF guidelines state that BMD testing is indicated in the following individuals: • Women age 65 and older and men age 70 and older, regardless of clinical risk factors • Younger postmenopausal women, women in the menopausal transition and men age 50 to 69 with clinical risk factors for fracture • Adults who have a fracture after age 50 • Adults with a condition (e.g., rheumatoid arthritis) or taking a medication (e.g., glucocorticoids in a daily dose ≥ 5 mg prednisone or equivalent for ≥ three months) associated with low bone mass or bone loss	HTAS is aware of the NOF guideline. Methodology for production of the guideline is not described. Funding of the NOF includes a substantial number of industry donors, including Pfizer, Medtronic, Novartis and 15 others.
	3	The NOF guidelines also describe the use of DXA to monitor osteoporosis therapy, as follows: • Serial central DXA testing is an important component of osteoporosis management. • Measurements for monitoring patients should be performed in accordance with medical necessity, expected response and in consideration of local regulatory requirements. NOF recommends that repeat BMD assessments generally agree with Medicare guidelines of every two years, but recognizes that testing more frequently may be warranted in certain clinical situations.	See comment #2. There is no discussion in the NOF guideline about test characteristics (i.e., precision) of DXA; retesting too soon may result in the margin of error of the test being larger than the actual change in the value of the bone density. USPSTF recommendation states: "Because of limitations in the precision of testing, a minimum of 2 years may be needed to reliably measure a change in BMD;

DRAFT

Expert	#	Comment	Disposition
		Clinical situations for which testing more frequently (e.g., one year interval) is helpful includes patients started on treatment or changing treatment in order to evaluate for treatment effect, and patients on glucocorticoid therapy who are at risk for rapid bone loss.	however, longer intervals may be necessary to improve fracture risk prediction." Current coverage recommendations allow for more frequent testing in patients for whom there has been a significant change in risk factors other than medication therapy.
	4	Although concerns have been raised that some screening prevention programs for other chronic diseases do not result in healthcare savings (12), this is not the case for BMD testing in appropriately selected patients. The experience of healthcare systems suggests that increases in BMD testing reduce fracture rates and save money. A 5-year observational study evaluated the clinical and fiscal outcomes of the Geisinger Health System Osteoporosis Disease Management Program from 1996 to 2000 (13). It was found that implementation of osteoporosis guidelines that included increases in BMD testing and treatment was associated with a significant decrease in the age-adjusted incidence of hip fractures and an estimated $7.8 million reduction in healthcare costs during this 5-year period.	This observational study projected cost savings of this screening program in women over 65, but projected additional expense in the population between 55 and 65. Guidance document recommends screening on all women 65 and over.
	5	At Kaiser Southern California, an osteoporosis disease management program ("Healthy Bones Program") was fully implemented in 2002, with a goal of reducing hip fractures by increasing BMD testing rates and treatment in patients at high risk of hip fracture (14;15). It was estimated that in 2006, 935 hip fractures, with an average cost of $33,000 each, were prevented, resulting in savings of over $30.8 million for Kaiser (16). Multiple osteoporosis screening strategies have been found to be clinically effective and cost-effective as well (17-19).	Ref #14 not available through OHSU library. Ref #15 is a clinical summary article that includes a brief description of Ref #16, which is a prospective observational study of the "Healthy Bones" program. This included screening of all women over 65, men over 70, patients with history of hip or fragility fracture or on steroids. Ref #17 is a CEA of a variety of different screening strategies. While they report the best strategy with ICER < $50,000 was initiation of screening at age 55 with DXA and rescreening every 5 years, they note that several strategies using SCORE (a screening tool similar to FRAX) for prescreening were more cost effective, with ICERs < $30,000. Ref #18 is a position statement of the American College of Preventive Medicine, which states: "All adult patients age ≥ 50 years should be evaluated for risk factors for osteoporosis. Screening with BMD testing for osteoporosis is recommended in women aged 65 years and in men aged 70 years. Younger postmenopausal women and men aged 50–69 years should undergo screening if they have at least one major or two minor risk factors for osteoporosis." Ref #19 is also a CEA

DRAFT

Expert	#	Comment	Disposition
			that concludes "bone densitometry of post-menopausal women who have not had a prior fracture is reasonable from 65-70 years of age, and is perhaps reasonable for men without a prior fracture after the age of 80 years depending on drug costs, the direct medical costs of fractures, fracture disutility, underlying fracture rates in the population and the societal willingness to pay for health benefits."
	6	**Comments on HERC coverage guidance.** Three sources of medical evidence were used in the development of the coverage guidance: 1. USPSTF recommendations for screening for osteoporosis (20;21); 2. a posthoc subgroup analysis of a single observational study in postmenopausal women (22); and the NICE guidelines from the UK (23). There are serious concerns with each of these that limit their applicability in setting rules for DXA coverage in the US.	HTAS acknowledges that these are the source documents, but disagrees that there are serious concerns regarding their use.
	7	**USPSTF recommendations-** The USPSTF recommended screening for osteoporosis "in women aged 65 years or older and in younger women whose fracture risk is equal to or greater than that of a 65-year-old white woman who has no additional risk factors." This was taken almost verbatim for inclusion in the HERC Guidance. However, the proposal very difficult to implement in clinical practice, as it would involve using FRAX without the benefit of BMD, which is not as good a predictor of fracture risk as FRAX with BMD, and assumes that physicians have the time and knowledge to use FRAX regularly and correctly. A 65 year-old Caucasian woman of average height and weight with no risk factors has a FRAX 10-year probability of major osteoporotic fracture of 9.4% and a 10-year probability of hip fracture of 1.4%. If she has low body weight, the numbers are 11% and 3.0%, respectively. If she is Hispanic, it is 6.0% and 1.7%, respectively. If she is Asian, it is 5.9% and 1.7%, respectively. If she is Black, it is 4.7% and 1.3%, respectively. If another fracture risk calculator, such as Garvan, is used for a 65-year old Caucasian woman with no risk factors, there is a 1.2% 5-year risk of hip fracture, a 2.4% 10-year risk of hip fracture, a 6.7% 5-year risk of any fragility fracture, and a 13.9% 10-year risk of any fragility fracture. There are other calculators as well that would generate different numbers. It is simply not feasible in a busy medical practice for any physician to sort through all of this and not possible for a regulatory agency to monitor for compliance.	The USPSTF selected the FRAX tool because "this tool relies on easily obtainable clinical information, such as age, body mass index (BMI), parental fracture history, and tobacco and alcohol use; its development was supported by a broad international collaboration and extensively validated in 2 large U.S. cohorts; and it is freely accessible to clinicians and the public." HTAS does not agree that it is not feasible for a physician to utilize this tool and believes that there are many who do. Compliance is an issue of implementation and does not impact the recommendations.
	8	The USPSTF addressed only screening DXA in women; they do not provide guidance on the use of DXA other than screening (e.g., monitoring) or DXA in men. It should be noted that	HTAS is aware that the USPSTF does not address the use of DXA in monitoring, and therefore includes the Gourlay

DRAFT

Expert	#	Comment	Disposition
		men age 70 and older are at high risk for fracture, and the consequences of fractures in men (morbidity and mortality) are more grave than in women. The adoption of the USPSTF recommendation would serve to reduce the use of DXA in evaluating patients (especially postmenopausal women under age 65 and men) for fracture risk, when the current problem is quite the opposite- too few patients are being screened for osteoporosis.	study in the guidance document to address this void. The USPSTF does address the use of DXA in men, stating that the evidence is insufficient to recommend for or against screening. .
	9	**Gourlay et al study-** This analysis of a subset of subjects in the Study of Osteoporotic Fractures (SOF) concluded quite reasonably that older women with very good BMD were unlikely to develop osteoporosis for many years, if ever. However, it was widely misinterpreted in the media, and by some healthcare providers, to mean that DXA is an expensive overused technology that was increasing medical expenses with little benefit. There was a firestorm of protest from many physicians and professional societies to set the story straight, including two where I was an author (24;25). Gourlay et al correctly identified limitations of the study that preclude its applicability to a wider patient population. The study cohort was restricted to pre-selected women ≥ 67 years of age and did not include men or younger postmenopausal women. It is particularly important to note that women in their early postmenopausal years are likely to experience accelerated bone loss that may require short testing intervals (e.g., 1-2 years) to assess. Also excluded from the trial were nearly 50% of the SOF study participants who had a previous diagnosis of osteoporosis (based on a prior hip or clinical vertebral fracture or densitometric evidence of osteoporosis) or who were already on treatment for osteoporosis.	HTAS is aware of the limitations of the Gourlay study. However, no other evidence has been identified or provided that provides evidence supporting a different testing interval. The cited reference #24 is an editorial that is verbatim to the comment provided here. Reference #25 is a letter to the editor. The author's (Gourlay's) response is as follows: "We strongly agree with Lewiecki and colleagues that too few initial BMD tests are performed in older women. An appropriate response to our results would be for primary care physicians to substantially increase the number of initial tests in older women, then to tailor the subsequent BMD screening interval according to BMD T-score and age."
	10	There were other limitations not noted by the authors. Only clinical vertebral fractures were considered in the analysis, although undiagnosed morphometric vertebral fractures are common in patients with densitometric evidence of osteopenia and are associated with high morbidity (26).	Ref #26 is a prospective case series that followed women > 65 over 4 years and reported incidence of vertebral fracture and back pain/disability. It found that approx. 2/3 of new fractures were not diagnosed clinically, yet those patients still reported increased pain and disability. These fractures were diagnosed by lateral spine radiographs, which would not be indicated in the general population. Unclear how this relates to the recommended guidance, or how this suggests the need for more frequent monitoring.
	11	In a prospective cohort study of 671 postmenopausal women undergoing periodic spine imaging, 48% of vertebral fractures were found in women with T-scores between -1.0 and -2.5. With a morphometric vertebral fracture, they would be reclassified as having a clinical diagnosis of osteoporosis (27). Many of these patients would not have been identified in the study of Gourlay et al.	Ref #27 is a prospective case series of 671 post-menopausal women followed over 9 years. This study found that women who were osteopenic had an increased risk of fracture over that time period, and risk was also increased with age, prior fracture and high bone turnover

DRAFT

Expert	#	Comment	Disposition
			markers. There is no comment in the article regarding reclassification of these women as having osteoporosis. WHO criteria and NOF guideline list only T-score as criteria.
	12	In making treatment decisions in clinical practice, it is imperative to consider risk factors for fracture in addition to the femoral neck and total hip T-score. Gourlay et al., for example, did not measure lumbar spine BMD. Low lumbar spine BMD is associated with increased fracture risk at all skeletal sites (28). Moreover, lumbar spine T-score may be ≤ -2.5 even if the femoral neck or total hip T-score is > -2.5. Without tracking lumbar spine BMD, Gourley et al. may have underestimated the number of individuals who progressed to osteoporosis during the study. Most importantly, with its singular focus on BMD, the study did not capture those patients with osteopenia who by the FRAX fracture risk assessment would have been at high risk for fracture and therefore warrant drug therapy. It would be grossly inappropriate to use the Gourlay et al study to set guidelines for frequency of BMD testing in the vast majority of clinical practice patients.	The abstract of Ref #28 states this was a prospective case series of 8,134 women > 65 followed 0.7 years and found the risk of fracture inversely related to BMD at all sites of measurement (proximal femur, spine, calcaneus, distal radius, proximal radius), and that none were more predictive than others. Does not appear to support contention that spine BMD needs to be tracked in addition to or instead of hip BMD. While the Gourlay article only evaluated BMD, again, no other evidence has been identified or provided that provides evidence supporting a different testing interval.
	13	**NICE guidelines-** These guidelines were developed through economic modeling of circumstances in the UK, where healthcare priorities and resources are quite different than in the US. This modeling used economic assumptions, including fracture-related medical expenses, that are uncertain even in the UK, and clearly not applicable in the US. FRAX in the UK was calibrated using country-specific fracture prevalence rates and mortality statistics that are not the same as in the US. There is controversy regarding the NICE guidelines amongst healthcare providers in the UK. As with all guidelines, NICE recognize that healthcare decisions should be individualized according the needs each patient.	HTAS does not disagree that modeling and economic assumptions in the UK may not apply perfectly to the US setting, but evidence to support an alternative testing schedule has not been provided. HTAS is familiar with controversy over testing guidelines, and while it is ideal for healthcare decisions to be individualized, that does not eliminate the need for a population-based coverage decision.
	14	**Recommendations.** It is my opinion that the proposed HERC Coverage Guidance, while well intentioned, is not sufficiently clear for clinical use, and that it would not be in the best interests of the citizens of Oregon to implement as it is. I think Oregon could do no better than to adopt the NOF guidelines for BMD testing and frequency of testing, allowing for physicians to individualize patient care decisions as needed. There are a number of minor formatting issues that should be corrected according to standard nomenclature established by the International Society for Clinical Densitometry (29). Change "DEXA" to "DXA," which is the preferred acronym. Be consistent in using "T-score" and not other forms, such as "T score," and express T-scores to one decimal place not two. Note that "advanced osteopenia" is not a recognized diagnostic category and should not be used; it was presented by the authors of the Gourlay et al study for use in their publication but has no established definition.	Some formatting corrections have been made, thank you. The use of 2 decimal points has been preserved, as this is directly from the evidence source. "Advanced osteopenia" is not deleted, as it is a helpful description of the T-score value 2.0 to 2.49. HTAS does not believe the NOF guidelines are sufficiently evidence-based for adoption.

DRAFT

References Submitted by Commenter

1) Klibanski A, Adams-Campbell L, Bassford T, Blair SN, Boden SD, Dickersin K, et al. Osteoporosis prevention, diagnosis, and therapy. JAMA 2001 Feb 14;285(6):785-95.
2) US Department of Health and Human Services. Bone Health and Osteoporosis: A Report of the Surgeon General. Rockville, MD: US Department of Health and Human Services, Office of the Surgeon General; 2004.
3) Lewiecki EM. Update on bone density testing. Curr Osteoporos Rep 2005 Dec;3(4):136-42.
4) World Health Organization. FRAX WHO Fracture Risk Assessment Tool. World Health Organization 2010 [cited 2010 Oct 1];Available from: URL: http://www.shef.ac.uk/FRAX/
5) MacLean C, Newberry S, Maglione M, McMahon M, Ranganath V, Suttorp M, et al. Systematic review: comparative effectiveness of treatments to prevent fractures in men and women with low bone density or osteoporosis. Ann Intern Med 2008 Feb 5;148(3):197-213.
6) Delmas PD, van de Langerijt L, Watts NB, Eastell R, Genant H, Grauer A, et al. Underdiagnosis of vertebral fractures is a worldwide problem: the IMPACT study. J Bone Miner Res 2005 Apr;20(4):557-63.
7) Panneman MJ, Lips P, Sen SS, Herings RM. Undertreatment with anti-osteoporotic drugs after hospitalization for fracture. Osteoporos Int 2004 Feb;15(2):120-4.
8) Cramer JA, Gold DT, Silverman SL, Lewiecki EM. A systematic review of persistence and compliance with bisphosphonates for osteoporosis. Osteoporos Int 2007 Aug;18(8):1023-31.
9) Lewiecki EM. Review of guidelines for bone mineral density testing and treatment of osteoporosis. Curr Osteoporos Rep 2005 Sep;3(3):75-83.
10) Lewiecki EM. Bone densitometry and vertebral fracture assessment. Curr Osteoporos Rep 2010 Sep;8(3):123-30.
11) National Osteoporosis Foundation. Clinician's Guide to Prevention and Treatment of Osteoporosis. National Osteoporosis Foundation 2013 [cited 2013 Jun 14];Available from: URL: http://www.nof.org/files/nof/public/content/resource/913/files/580.pdf
12) Esserman L, Shieh Y, Thompson I. Rethinking screening for breast cancer and prostate cancer. JAMA 2009 Oct 21;302(15):1685-92.
13) Newman ED, Ayoub WT, Starkey RH, Diehl JM, Wood GC. Osteoporosis disease management in a rural health care population: hip fracture reduction and reduced costs in postmenopausal women after 5 years. Osteoporos Int 2003 Feb;14(2):146-51.
14) Dell R, Greene D. Is osteoporosis disease management cost effective? Curr Osteoporos Rep 2010 Mar;8(1):49-55.
15) Dell RM, Greene D, Anderson D, Williams K. Osteoporosis disease management: What every orthopaedic surgeon should know. J Bone Joint Surg Am 2009 Nov;91 Suppl 6:79-86.
16) Dell R, Greene D, Schelkun SR, Williams K. Osteoporosis disease management: the role of the orthopaedic surgeon. J Bone Joint Surg Am 2008 Nov;90 Suppl 4:188-94.
17) Nayak S, Roberts MS, Greenspan SL. Cost-effectiveness of different screening strategies for osteoporosis in postmenopausal women. Ann Intern Med 2011 Dec 6;155(11):751-61.
18) Lim LS, Hoeksema LJ, Sherin K. Screening for osteoporosis in the adult U.S. population: ACPM position statement on preventive practice. Am J Prev Med 2009 Apr;36(4):366-75.
19) Schousboe JT. Cost effectiveness of screen-and-treat strategies for low bone mineral density: how do we screen, who do we screen and who do we treat? Appl Health Econ Health Policy 2008;6(1):1-18.
20) U.S.Preventive Services Task Force. Screening for Osteoporosis: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med 2011 Jan 17;154:356-64.
21) Nelson HD, Haney EM, Dana T, Bougatsos C, Chou R. Screening for osteoporosis: an update for the U.S. Preventive Services Task Force. Ann Intern Med 2010 Jul 20;153(2):99-111.
22) Gourlay ML, Fine JP, Preisser JS, May RC, Li C, Lui LY, et al. Bone-density testing interval and transition to osteoporosis in older women. N Engl J Med 2012 Jan 19;366(3):225-33.

DRAFT

23) National Clinical Guideline Centre. Osteoporosis: fragility fracture risk. National Clinical Guideline Centre 2012 [cited 2013 Jun 14];Available from: URL: http://www.nice.org.uk/nicemedia/live/13857/60400/60400.pdf
24) Lewiecki EM, Laster AJ, Miller PD, Bilezikian JP. More bone density testing is needed, not less. J Bone Miner Res 2012 Mar 1.
25) Lewiecki EM, Miller PD, Bilezikian JP. Bone-density testing interval and transition to osteoporosis. N Engl J Med 2012 Apr 19;366(16):1546-7.
26) Nevitt MC, Ettinger B, Black DM, Stone K, Jamal SA, Ensrud K, et al. The association of radiographically detected vertebral fractures with back pain and function: A prospective study. Ann Intern Med 1998 May 15;128(10):793-800.
27) Sornay-Rendu E, Munoz F, Garnero P, Duboeuf F, Delmas PD. Identification of osteopenic women at high risk of fracture: the OFELY study. J Bone Miner Res 2005 Oct;20(10):1813-9.
28) Black DM, Cummings SR, Genant HK, Nevitt MC, Palermo L, Browner W. Axial and appendicular bone density predict fractures in older women. J Bone Miner Res 1992 Jun;7(6):633-8.
29) Baim S, Binkley N, Bilezikian JP, Kendler DL, Hans DB, Lewiecki EM, et al. Official Positions of the International Society for Clinical Densitometry and executive summary of the 2007 ISCD Position Development Conference. J Clin Densitom 2008 Jan;11(1):75-91.

DRAFT

HEALTH EVIDENCE REVIEW COMMISSION (HERC)

DRAFT COVERAGE GUIDANCE: TREATMENT OF SLEEP APNEA IN ADULTS

DRAFT for HTAS Meeting Materials 11/25/2013

HERC COVERAGE GUIDANCE

Coverage of treatment for Obstructive Sleep Apnea (OSA) in adults should be limited, as follows:

CPAP is recommended for coverage initially when all of the following conditions are met *(strong recommendation):*

- 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 30 events per hour; or if between 15 and 30 events with additional symptoms including one or more of the following:
  - excessive daytime sleepiness (Epworth Sleepiness Scale score > 10), or
  - documented hypertension, or
  - ischemic heart disease, or
  - 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 should be 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 recommended for coverage.

Intensive weight loss programs (if provided in the benefit package) are recommended for coverage for patients with obesity and obstructive sleep apnea.

Surgery for sleep apnea for adults is not recommended for coverage *(weak recommendation)*.

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 Heath 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

Gleitsmann, K., Kriz, H., Thielke, A., Bunker, K., Ryan, K., Lorish, K., & King, V. (2012). *Sleep apnea diagnosis and treatment in adults.* Produced for the Washington HTA Program. Olympia, WA: Center for Evidence-based Policy, Oregon Health and Science University for the Washington Health Technology Assessment Program. Retrieved September 13, 2012, from http://www.hta.hca.wa.gov/documents/sleep_apnea_final_report.pdf

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

SUMMARY OF EVIDENCE

Clinical Background

Obstructive sleep apnea (OSA) refers to sleep-disordered breathing due to the recurrent collapse of pharyngeal tissues resulting in snoring, fitful sleep, and daytime somnolence. These episodes are characterized by either reduced airflow (hypopnea), or a complete obstruction (apnea), with a subsequent drop in oxygen saturation, interfering with gas exchange. Obstructive sleep apnea is a cause of significant morbidity and mortality and is associated with hypertension, neuropsychological impairment, motor vehicle accidents, stroke, cardiovascular disease, diabetes, and decreased quality of life. The prevalence of OSA is 2 to 7% in the general adult population. Prevalence increases steadily with age, to approximately 20% among people older than age 60. Risk factors for OSA include male gender, age, obesity, airway characteristics, familial/genetic predisposition, smoking, and alcohol consumption. The majority of patients with OSA are asymptomatic, unaware of their sleep disordered breathing and associated health risks.

The diagnosis as well as the treatment of OSA is complicated by the difficulty in defining the syndrome. There is controversy surrounding the parameters to be used in a clinical definition as well as which diagnostic method is most appropriate to detect OSA. The

DRAFT

current standard for diagnosing OSA is polysomnography (PSG) administered in a sleep study facility. The frequency of obstructed breathing events (i.e., the apnea-hypopnea index (AHI)), combined with multiple other clinical features of obstruction (e.g., oxygen desaturation, air flow, choking episodes) are recorded during sleep. A diagnosis of OSA is generally made when AHI is greater than or equal to 15 or greater than 5 with noticeable daytime symptoms.

When considering the diagnosis of sleep apnea and the relationship between apnea/hypopnea index (AHI) and long term outcomes, the WA HTA report limited inclusion criteria to longitudinal studies of at least 500 participants and a minimum of 1 year of follow up. Eleven trials were included in total. Four evaluated AHI as a predictor of mortality, and of those, three evaluated AHI categories (mild, moderate, severe). All found that AHI > 30 had a significant increased risk of death compared to AHI < 5-10. Those with AHI between 10 and 30 had a non-significantly increased risk of death.

Other conditions for which a correlation with AHI has been examined include non-fatal cardiovascular disease, stroke, diabetes and hypertension. There was a significant positive correlation between AHI of > 30 and non-fatal cardiovascular disease in patients not treated with CPAP. A similar correlation was not seen for lower levels of AHI. For stroke, there was no overall increase in incident stroke over 12 years of follow up in patients with AHI > 20. For incident hypertension, results were mixed. One study found that AHI was not an independent predictor of incident hypertension unless BMI was not controlled for in the analysis. The other study found a significant association between any AHI > 0 and the presence of hypertension at 4 and 8 years follow up, with higher AHI having a stronger association. For type 2 diabetes, results were again mixed. One study found no association between AHI and the incidence of diabetes after four years, while another found a significant association after 2.7 years for AHI > 8, There was no association between baseline AHI and quality of life (QOL) in the one study that reported on it after 5 years.

There have been various modalities developed to treat OSA, most attempting to reduce the airway obstructive component. Continuous positive airway pressure (CPAP) is the first-line therapy for OSA and opens the airway with compressed air. However, the CPAP machinery required is poorly tolerated and compliance is a major concern. Various oral appliances, which attempt to splint open the airway, have been used as an alternative to CPAP. Surgical procedures, including various surgeries on the oropharyngeal anatomy to alter airway mechanics, are performed to treat OSA. Bariatric surgery may be performed to reduce the volume of obstructive tissues. Other interventions that have been used to treat OSA include: weight loss regimens; smoking cessation; caffeine and alcohol avoidance; positional therapy; oropharyngeal physical therapy to strengthen the musculature and reduce obstruction; arrhythmia treatment for

nocturnal bradycardia; complementary and alternative medicine (e.g., acupuncture), and a variety of pharmacologic agents.

Evidence Review

Continuous Positive Airway Pressure

A moderate strength of evidence was found for the effectiveness of treatment of OSA with CPAP. However, there was insufficient evidence to determine which patients CPAP might benefit the most. When evaluating the effectiveness of CPAP, 22 trials were included that had a range of baseline AHI from 10 to 65. With regard to inclusion criteria:

- 9 required AHI >5
- 1 required AHI > 10
- 7 required AHI > 15
- 2 required AHI > 20
- 1 required AHI > 30
- 2 did not report baseline or required AHI

Only one of these evaluated an objective clinical outcome, and it found no significant effect of CPAP on CHF symptoms (baseline average AHI 27). When evaluating the Epworth Sleepiness Scale[1] (ESS) as an outcome, a total of 14 trials were included. Of the seven that included patients with baseline AHI as low as 5, only three found a statistically significant benefit of CPAP on ESS. Of those three, only one had an average baseline AHI for the study population less than 15. All of the studies that were limited to patients with an AHI of at least 15 found statistically significant benefit of CPAP. Improvements in ESS range from 2 to 7 points. Of the 3 trials that allowed AHI as low as 5 and found a significant difference, the improvements in ESS were 3 points (2 trials, average baseline AHI = 19 and 10) and 4 points (average baseline AHI = 27). A 1 point change in ESS is considered clinically significant.

Seven studies evaluated blood pressure; none found statistically significant differences between CPAP and control (minimum baseline AHI ranged from >5 to >30). One evaluated HbA1c and also found no difference (minimum baseline AHI >15). Ten studies reported on 29 different QOL measures. Overall, 11 measures in 6 trials reached statistical significance. Of those, only one had an average baseline AHI of less than 15 (range for remaining studies was 19 to 58).

The reviewed studies report sufficient evidence supporting large improvements in sleep measures with CPAP compared with control (e.g., reducing apnea hypopnea index

[1] A self-administered questionnaire that measures sleep propensity, total score ranges 0-24. Reference range is defined as ≤ 10, with 1 point change considered clinically significant. Sensitivity 49% and specificity 80% for detecting OSA using an AHI cutoff of 5 events/hour, based on one high quality study.

(AHI), improving symptoms as measured by the ESS, reducing arousal index, and raising the minimum oxygen saturation). Weak evidence demonstrated no consistent benefit in improving quality of life, neurocognitive measures or other intermediate outcomes.

Despite no or weak evidence for an effect of CPAP on clinical outcomes, given the large magnitude of effect on the intermediate outcomes of AHI and ESS, the strength of evidence that CPAP is an effective treatment to alleviate sleep apnea signs and symptoms was rated moderate. However, the link between AHI reduction and long term clinical outcomes is not directly proven. There was insufficient evidence regarding most comparisons of various different CPAP devices, including nasal vs. oral, bilevel vs. fixed, flexible bilevel vs. fixed and humidified vs. non-humidified. However, there was a low strength of evidence that C-Flex (a proprietary CPAP technology that reduces the pressure slightly at the beginning of exhalation) is not significantly different than fixed CPAP in compliance or other outcomes, and a moderate strength of evidence that autoCPAP and fixed CPAP result in similar compliance and treatment effects.

Other Treatments for Obstructive Sleep Apnea

Mandibular advancement devices (oral appliances) had moderate strength of evidence supporting their use as an effective treatment for OSA. However, as with CPAP, there was insufficient evidence to indicate which patients might benefit from their use. There was moderate evidence that the use of CPAP is superior to mandibular advancement devices with regard to improved sleep study measures, but weak evidence that there is minimal difference between the two for improving compliance, treatment response, quality of life, or neurocognitive measures. There was insufficient evidence to compare the different oral devices, other than mandibular advancement devices.

Six surgical interventions for the treatment of OSA were reviewed (uvulopalatopharyngoplasty [UPPP], laser-assisted uvulopalatoplasty [LAUP], radiofrequency ablation [RFA], and combinations of pharyngoplasty, tonsillectomy, adenoidectomy, genioglossal advancement septoplasty, radiofrequency ablation of the inferior nasal turbinates, or combination nasal surgery) compared to sham, conservative therapy or no treatment. No surgical interventions were compared to each other. Details of each study are presented below:

Back 2009 compared a single session of RFA surgery of the soft palate to sham surgery (simulated surgery with no energy administered). The study included 32 male patients with mild sleep apnea (AHI 5-15 events/hr) and habitual snoring following a failed trial of conservative treatment (weight loss, positional therapy, restriction of alcohol and sedatives). At 4 month followup, no statistically significant difference between groups in AHI, ESS, minimum oxygen saturation, and quality of life [as measured by the Short Form 36 questionnaire (SF-36)] were found.

Koutsourelakis 2008 randomized patients to either nasal surgery (submucous resection of the deviated septum and bilateral resection of inferior turbinates) or sham surgery (simulated nasal surgery under anesthesia). In addition to OSA (defined as AHI ≥5 events/hr), all patients had fixed nasal obstruction due to deviated nasal septum. The study was conducted on 49, predominately male patients with a mean baseline AHI of 31 events/hr. After 4 months followup, the study found no statistically significant difference between groups in AHI or on ESS.

Woodson 2003 conducted a three-arm RCT that included a comparison of multilevel temperature controlled RFA of the soft palate with sham surgery (simulated RFA with no energy delivered). The study was conducted in 51, predominately male patients. Notably, the age of participants between groups was significantly different at baseline. (49 years (RFA) versus 51 years (sham), P=0.04). The mean baseline AHI also differed among groups (21 (RFA) versus 15 (sham) events/hr; P=0.06, including the CPAP study group). After 8 weeks followup, the study found a significantly greater improvement in sleep quality as measured by Functional Outcomes of Sleep Questionnaire with RFA as compared to sham surgery (P=0.04), but no statistically significant difference in AHI, ESS, minimum oxygen saturation, or quality of life as measured by SF-36.

Ferguson 2003 randomized patients to either LAUP or no treatment. In LAUP, the uvula and a specified portion of the palate is vaporized under local anesthesia in an outpatient setting. The goal is to relieve obstruction in patients with mild OSA or snoring. The study included 44 mostly male patients with mild OSA (AHI 10-27 events/hr) and snoring. This study reported disparate followup durations of 15 months in the LAUP group and 8 months in the control group. A statistically significant improvement in AHI was observed following LAUP as compared with no treatment (net change -10.5 events/hr; P=0.04). However, there was no statistically significant difference between groups on the ESS or in quality of life as measured by Sleep Apnea Quality of Life Index.

Guilleminault 2008 was reported as a crossover study comparing several surgical combinations to cognitive behavioral therapy in 30 patients with insomnia and mild OSA (mean AHI 10 events/hr). Based on anatomy, disease severity, and comorbidity, patients received combinations of pharyngoplasty, tonsillectomy, adenoidectomy, genioglossal advancement septoplasty, and RFA of the inferior nasal turbinates. Only the first phase of the trial was evaluated. Results showed that surgery led to improvements in AHI (-6.2 events/hr; P=0.0001), ESS (-1.1; P=0.002), minimum oxygen saturation (4.4 percent; P=0.0001) and two other sleep measures as compared to cognitive behavioral therapy.

Lojander 1996 & 1999 compared UPPP with or without mandibular osteotomy to conservative treatment (weight loss, positional therapy, and avoidance of tranquilizers

DRAFT

and alcohol at bedtime). The study included 32, predominately male patients with a mean age of 47 years and a mean baseline BMI of 31 kg/m2. Baseline Oxygen Desaturation Index ranged from 10 to 72 events/hr. A significant improvement in daytime somnolence (net difference -25 on a visual analogue scale ranging from 0 (no somnolence) to 100 (worst); P<0.05) was observed after 12 months; no statistically significant difference was found between groups in cognitive function.

Li 2009, in a nonrandomized prospective study, compared correction of nasal septum and volume reduction of the inferior turbinates to conservative nasal treatments in patients with snoring, nasal obstruction, and OSA. The study included 66 patients, 44 of whom had surgery. The patients were almost all male, with a mean age of 38 years and a mean BMI of 26.2 kg/m2. Baseline AHI was 38 events/hr in the surgically treated group and 26 in the conservative treatment group (no significant difference), and baseline ESS was 10.6. The article did not report at what time point follow-up data were collected. The study found a statistically significant difference in ESS, favoring surgery (net difference -3.6; 95 percent CI -6.1, -1.1; P=0.02). The study found no difference in AHI, minimum oxygen saturation or two sleep measures.

Overall there was insufficient evidence with which to evaluate the efficacy of any of these surgical treatments. When each modality was compared to CPAP, the evidence was insufficient to determine their relative merits. No evidence that met inclusion criteria was identified for any other surgical procedures.

Of the other treatments for OSA that were considered, only intensive weight loss programs were an effective treatment in obese patients with OSA with a low strength of evidence. The remainder of the other management modalities (e.g., atrial overdrive pacing, medications, palatal implants, oropharyngeal exercises, tongue-retaining devices with positional alarms either in isolation or in combination, bariatric surgery, acupuncture, and auricular plaster) had insufficient evidence to determine the effects of using them for treatment of OSA.

Compliance with Treatment

Compliance in OSA patients prescribed nonsurgical treatments had moderate strength of evidence that compliance was greater with CPAP use with more severe OSA and insufficient evidence regarding potential predictors of mandibular advancement devices compliance.

The strength of evidence is low for indentifying any specific intervention which may improve CPAP compliance. No intervention type (e.g., education, telemonitoring) was more promising than others.

Overall Summary

CPAP is effective for improving sleep measures (e.g., reducing AHI, improving symptoms as measured by the Epworth Sleepiness Scale, reducing arousal index, and raising the minimum oxygen saturation), but there is no evidence of consistent benefit in improving quality of life, neurocognitive measures or other intermediate outcomes. There is more evidence for effectiveness in patients with higher (>15) AHI. AutoCPAP and fixed CPAP result in similar compliance and treatment effects. Mandibular advancement devices are effective treatment for OSA, although CPAP is superior to mandibular advancement devices with regard to improved sleep study measures. The evidence is insufficient to evaluate the efficacy of all surgical procedures and other treatments except intensive weight loss for obese patients with OSA.

[Evidence Source]

COMMITTEE DELIBERATIONS – HTAS

At the May 21, 2012 meeting, subcommittee members requested to add CMS criteria for CPAP compliance (70% of nights and 4 hours per night). Members requested further information to guide the decision about whether to perform surgery. At its June 25, 2012 meeting the subcommittee added language allowing coverage for surgery under certain conditions, and requested that the report be put out for public comment. On November 26, 2012 the subcommittee reviewed public comment and added a recommendation for coverage for intensive weight loss and the inclusion of the Epworth Sleepiness Scale score > 10 as a requirement for a CPAP trial. It removed the reference to impaired cognition before referring the draft coverage guidance to HERC.

COMMITTEE DELIBERATIONS – VBBS

At its March 14, 2013 meeting, the Value-based Benefits Subcommittee discussed the draft coverage guidance and recommended changing it in order to allow coverage for surgery only after both CPAP and an oral appliance had failed.

HERC DELIBERATIONS

In its review May 9, 2013, the HERC requested that staff consider the evidence around coverage for surgeries, creating a GRADE-informed framework and HERC Guidance Development Framework for this service, as has been done for the newer coverage guidances. These have been added as Appendices A, B and C. They asked that if the recommendation comes down as “not recommended for coverage” that the coverage guidance and associated coverage and prioritization decisions for the Oregon Health Plan, be referred back to VbBS without the coverage guidance returning to HTAS.

At its August 8, 2013 meeting, HERC reviewed additional evidence on the effectiveness of CPAP and returned the draft coverage guidance to the HTAS for additional work on surgery and indications for CPAP coverage, indicating that the document should go out for public comment again if changes are made which don't result from public comment.

COMMITTEE DELIBERATIONS – HTAS

At its September 23, 2013 meeting, based on the additional evidence reported, the HTAS changed the draft coverage guidance to recommend coverage for CPAP fo patients with AHI of at least 30, as well as for patients with specified symptoms and an AHI of at least 15. The subcommittee also changed its recommendation for surgery to a weak recommendation not to cover surgeries.

APPLICABLE CODES

CODES	DESCRIPTION
ICD-9 Diagnosis Codes	
327.20	Organic sleep apnea, unspecified
327.21	Primary central sleep apnea
327.23	Obstructive sleep apnea (adult) (pediatric)
327.27	Central sleep apnea in conditions classified elsewhere
327.29	Other organic sleep apnea
780.5	Sleep disturbance, unspecified
780.51	Insomnia with sleep apnea, unspecified
780.53	Hypersomnia with sleep apnea, unspecified
780.54	Hypersomnia, unspecified
780.57	Unspecified sleep apnea
ICD-9 Volume 3 (Procedure Codes)	
21.31	Nasal surgery (remove polyps)
21.88	Other septoplasty
27.64	Insertion of palatal implant
27.69	Uvulopalatopharyngoplasty
28.2	Tonsillectomy
28.3	Tonsillectomy/adenoidectomy
28.6	Adenoidectomy
31.29	Tracheostomy
93.9	CPAP
CPT Codes	
21198	Osteotomy, mandible
21199	Osteotomy, mandible, with genioglossus advancement
21206	Osteotomy, maxilla
21685	Hyoid myotomy and suspension
31600	Tracheostomy
41512	Tongue base suspension, permanent suture technique
41530	Radiofrequency reduction of the tongue base

CODES	DESCRIPTION
42145	Uvulopalatopharyngoplasty
42299	Unlisted procedure, palate, uvula (use for laser assisted uvulopalatoplasty (LAUP), somnoplasty, palatal implants)
HCPCS Codes	
A4604	Tubing with integrated heating element for use with positive airway pressure device
A7033	Pillow for use on nasal cannula type interface, replacement only, pair
A7034	Nasal interface (mask or cannula type) used with positive airway pressure device, with or without head strap
A7035	Headgear used with positive airway pressure device
A7036	Chinstrap used with positive airway pressure device
A7037	Tubing used with positive airway pressure device
A7038	Filter, disposable, used with positive airway pressure device
A7039	Filter, nondisposable, used with positive airway pressure device
A7524	Tracheostoma stent/stud/button, each
E0470	Respiratory assist device, bi-level pressure capability, without backup rate feature, used with noninvasive interface, e.g., nasal or facial mask (intermittent assist device with continuous positive airway pressure device)
E0471	Respiratory assist device, bi-level pressure capability, with back-up rate feature, used with noninvasive interface, e.g., nasal or facial mask (intermittent assist device with continuous positive airway pressure device)
E0472	Respiratory assist device, bi-level pressure capability, with backup rate feature, used with invasive interface, e.g., tracheostomy tube (intermittent assist device with continuous positive airway pressure device)
E0485	Oral device/appliance used to reduce upper airway collapsibility, adjustable or nonadjustable, prefabricated, includes fitting and adjustment
E0486	Oral device/appliance used to reduce upper airway collapsibility, adjustable or nonadjustable, custom fabricated, includes fitting and adjustment
E0601	Continuous airway pressure (CPAP) device

Note: Inclusion on this list does not guarantee coverage

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-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	Balance between desirable and undesirable effects	Quality of evidence*	Resource Allocation	Values and preferences	Coverage Recommendation
Surgery	Uncertain, but no certain benefit, and significant risk of surgery	Very low	Moderately costly	Moderate variability	Surgery for sleep apnea for adults is not recommended for coverage.
CPAP for patients with AHI 5-14 with symptoms/signs	No benefit on mortality or comorbid diseases (hypertension, diabetes, etc), minimal benefit on sleepiness/QOL, if any. No serious harms, but significant patient inconvenience.	Moderate[2]	Moderately costly	Moderate variability	CPAP coverage is not recommended at AHI levels less than 15.
CPAP for patients with AHI 15-29	No benefit on mortality or comorbid diseases (hypertension, diabetes, etc), moderate benefit on sleepiness/QOL. No serious harms, but significant patient	Moderate	Moderately costly	Moderate variability	CPAP coverage is recommended at AHI levels between 15 and 30 with daytime sleepiness, hypertension, ischemic

[2] The authors of the AHRQ report say, "Despite no or weak evidence for an effect of CPAP on clinical outcomes, given the large magnitude of effect on the intermediate outcomes of AHI and ESS, the strength of evidence that CPAP is an effective treatment to alleviate sleep apnea signs and symptoms was rated moderate."

Indication	Balance between desirable and undesirable effects	Quality of evidence*	Resource Allocation	Values and preferences	Coverage Recommendation
	inconvenience.				heart disease, or history of stroke.
CPAP for patients with AHI ≥ 30	Significant benefit on mortality/ comorbid diseases, moderate benefit on sleepiness/QOL. No serious harms, but significant patient inconvenience.	Moderate	Moderately costly	Small variability	CPAP coverage is recommended at AHI levels ≥ 30.

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

Note: GRADE framework elements are described in Appendix B

DRAFT

Appendix B. 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 of evidence across studies for the treatment/outcome

High = Further research is very unlikely to change our confidence in the estimate of effect.

Moderate = Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.

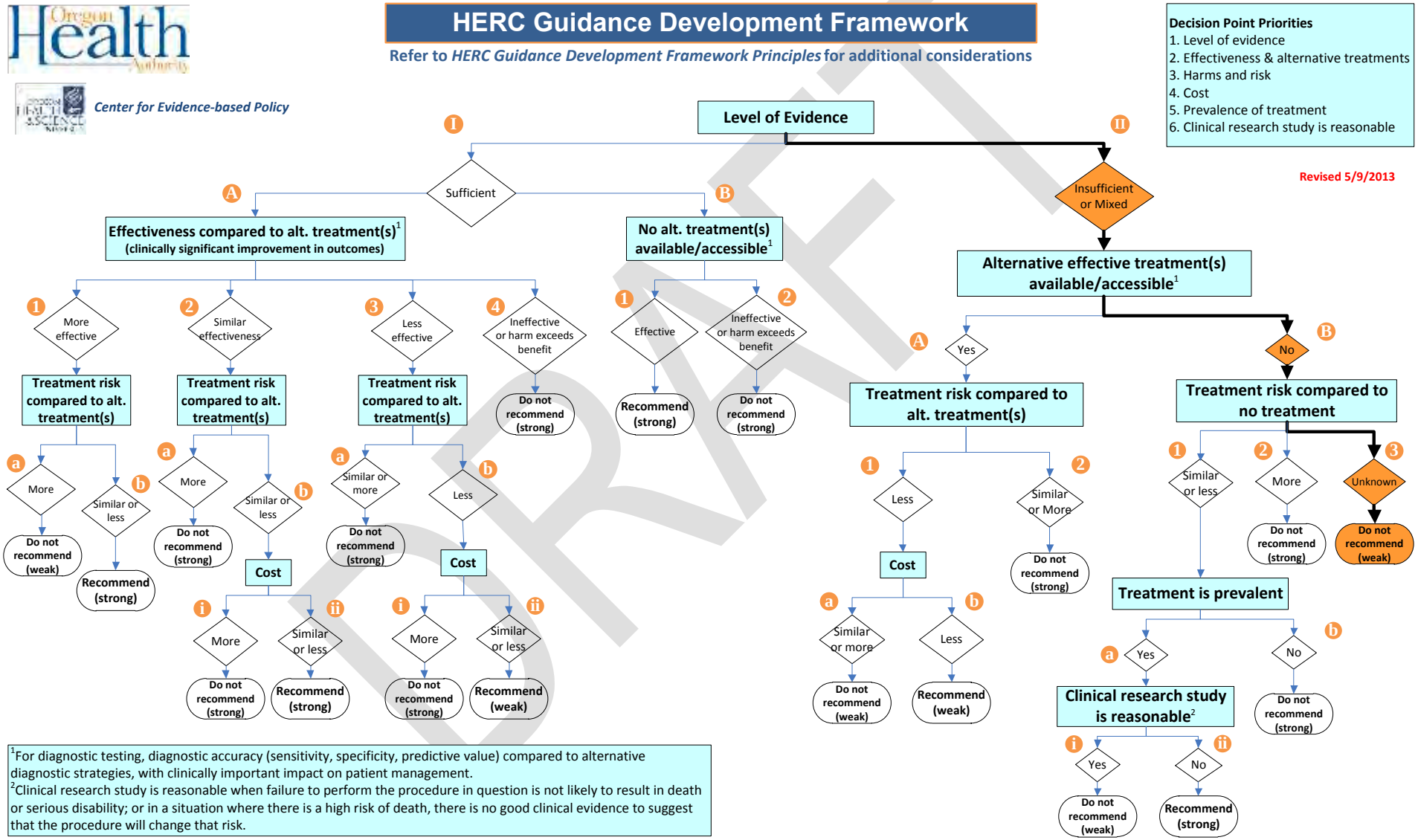
Low = Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.

Very low = Any estimate of effect is very uncertain.

DRAFT

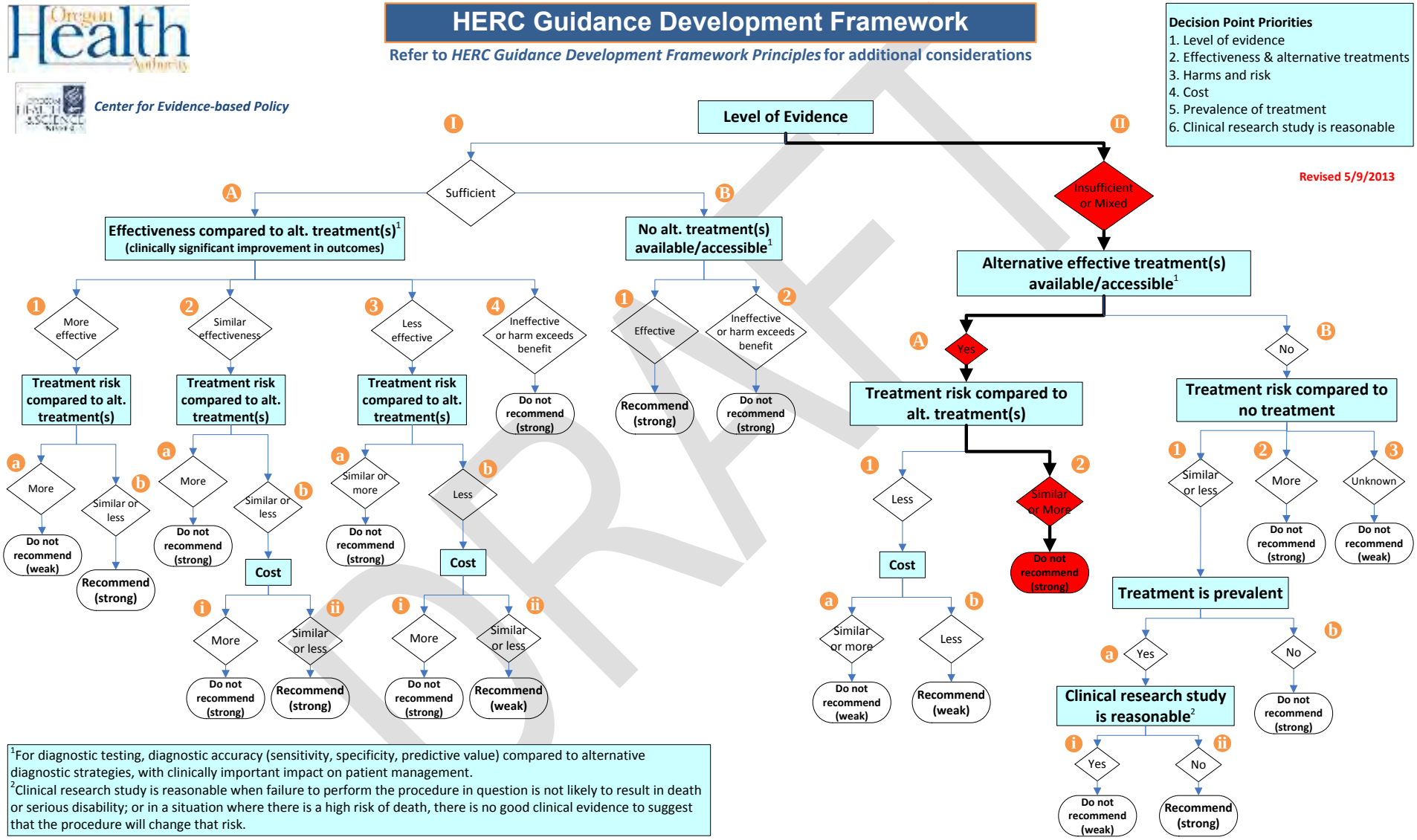
Appendix C. HERC Guidance Development Framework

Surgery for treatment of sleep apnea in adults when both CPAP and/or other alternatives (e.g., oral appliances) have failed

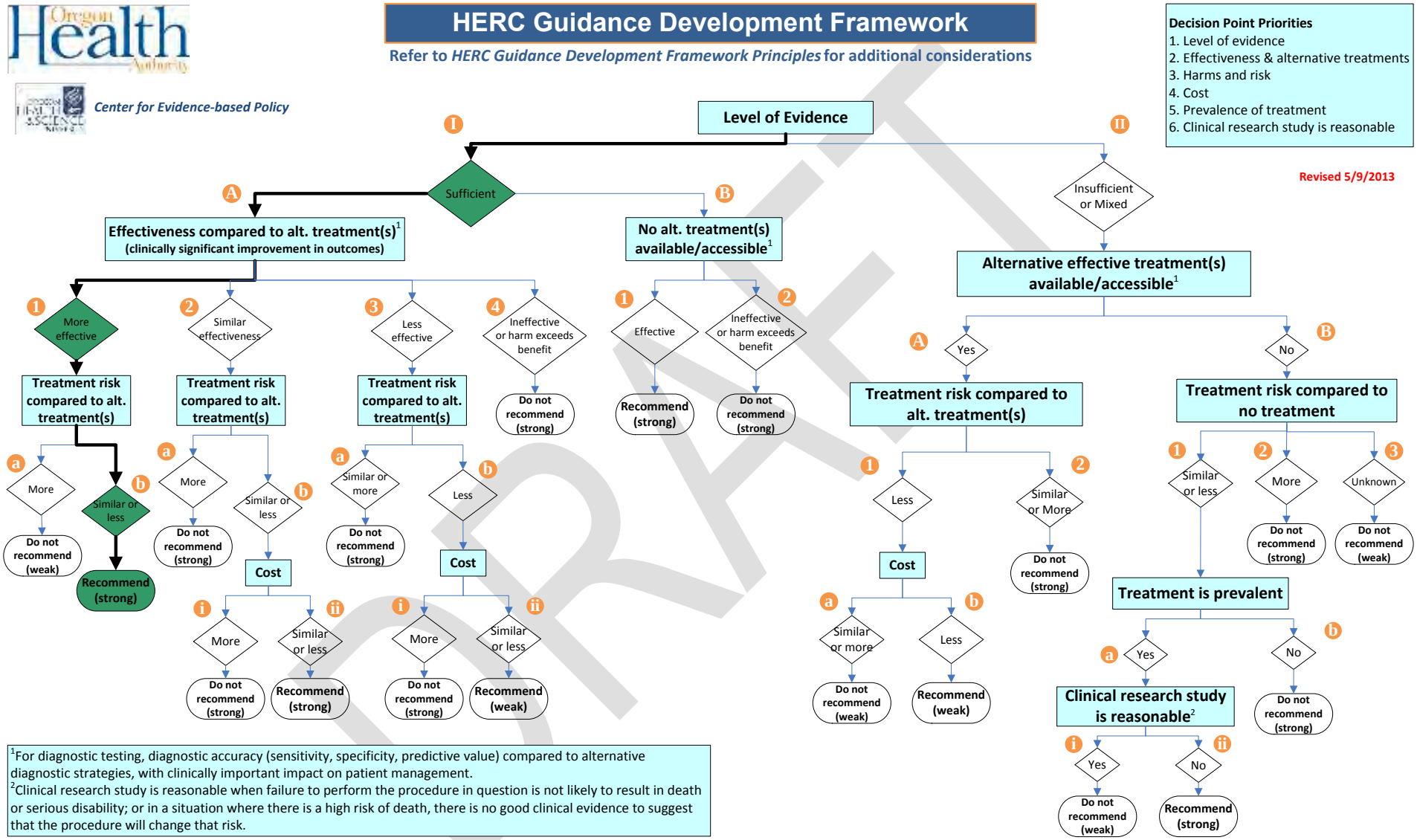


[1]For diagnostic testing, diagnostic accuracy (sensitivity, specificity, predictive value) compared to alternative diagnostic strategies, with clinically important impact on patient management.
[2]Clinical research study is reasonable when failure to perform the procedure in question is not likely to 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.

CPAP for Patients with AHI 5-14 with Symptoms/Signs (Compared to Oral Appliances)



CPAP for Patients with AHI 15-29; CPAP for Patients with AHI ≥ 30



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

DRAFT

Supplemental Review of the Washington HTA report in order to further inform the HERC Coverage Guidance on Treatment of Sleep Apnea

Source report: Gleitsmann, K., Kriz, H., Thielke, A., Bunker, K., Ryan, K., Lorish, K., & King, V. (2012). *Sleep apnea diagnosis and treatment in adults.* Produced for the Washington HTA Program. Olympia, WA: Center for Evidence-based Policy, Oregon Health and Science University for the Washington Health Technology Assessment Program

When considering the diagnosis of sleep apnea and the relationship between apnea/hypopnea index (AHI) and long term outcomes, the report limited inclusion criteria to longitudinal studies of at least 500 participants and a minimum of 1 year of follow up. Eleven trials were included in total. Four evaluated AHI as a predictor of mortality, and of those, three evaluated AHI categories (mild, moderate, severe). All found that AHI > 30 had a significant increased risk of death compared to AHI < 5-10. Those with AHI between 10 and 30 had a non-significantly increased risk of death.

Other conditions for which a correlation with AHI has been examined include non-fatal cardiovascular disease, stroke, diabetes and hypertension. There was a significant positive correlation between AHI of > 30 and non-fatal cardiovascular disease in patients not treated with CPAP. A similar correlation was not seen for lower levels of AHI. For stroke, there was no overall increase in incident stroke over 12 years of follow up in patients with AHI > 20. For incident hypertention, results were mixed. One study found that AHI was not an independent predictor of incident hypertention unless BMI was not controlled for in the analysis. The other study found a significant association between any AHI > 0 and the presence of hypertention at 4 and 8 years follow up, with higher AHI having a stronger association. For type 2 diabetes, results were again mixed. One study found no association between AHI and the incidence of diabetes after four years, while another found a significant association after 2.7 years for AHI > 8, There was no association between baseline AHI and quality of life (QOL) in the one study that reported on it after 5 years.

When evaluating the effectiveness of CPAP, a total of 22 trials were included that had a range of baseline AHI of 10 to 65:

- 9 had AHI >5
- 1 AHI > 10
- 7 AHI > 15
- 2 AHI > 20
- 1 AHI > 30
- 2 did not report baseline AHI

Only one of these evaluated an objective clinical outcome, and it found no significant effect of CPAP on CHF symptoms (baseline average AHI 27). When evaluating the Epworth Sleepiness Scale as an outcome, a total of 14 trials were included. Of the seven that included patients with baseline AHI as low as 5, only three found a statistically significant benefit of CPAP on ESS. Of those three, only one had an average baseline AHI for the study population less than 15. All of the studies that were limited to patients with an AHI of at least 15 found statistically significant benefit of CPAP. Improvements in ESS range from 2 to 7 points. Of the 3 trials that allowed AHI as low as 5 and found a significant difference, the improvements in ESS were 3 points (2 trials, average baseline AHI = 19 and 10) and 4 points (average baseline AHI = 27). A 1 point change in ESS is considered clinically significant.

Seven studies evaluated blood pressure; none found statistically significant differences between CPAP and control (minimum baseline AHI ranged from >5 to >30). One evaluated HbA1c and also found no difference (minimum baseline AHI >15).

Ten studies reported on 29 different QOL measures. Overall, 11 measures in 6 trials reached statistical significance. Of those, only one had an average baseline AHI of less than 15 (range for remaining studies was 19 to 58).

<u>Conclusion:</u> Degree of AHI and health outcomes

The only outcomes for which CPAP has proven benefit are improving Epworth sleepiness scale and a number of sleep study parameters, including AHI, arousal index, and oxygen saturation. Treatment with CPAP has not been proven to have an effect on morbidity (heart failure symptoms, blood pressure, HbA1c), although one trial that evaluated non-fatal cardiovascular events did find that the increased risk from an AHI > 30 was eliminated in patients using CPAP. There is some evidence of efficacy for improving quality of life, primarily for those with baseline AHI > 15. While AHI>30 is associated with increased mortality, there is no evidence that treatment with CPAP decreases that risk, although the longest study included in the evidence review was 12 months.

HERC Coverage Guidance – Treatment of Sleep Apnea in Adults
Disposition of Public Comments

General Comments

Stakeholder	#	Comment	Disposition
Medical Director, Health Plan Portland, OR	1	Regarding the Coverage Guidance, I have several suggestions for consideration. First would be to enhance the statement regarding excessive daytime sleepiness to require an objective evaluation of daytime sleepiness, presumably the Epworth Sleepiness Scale. This would avoid the subjectivity involved in any statement on the part of provider or DME supplier claiming member has "excessive sleepiness", without requirement of at least a standardized assessment. Likewise, "impaired cognition" is problematic in its subjectivity, although probably not wise to try and establish a standardized requirement for that condition, as it would likely lead to neuropsych testing requests, which would be of limited value in many cases (particularly if no baseline exists, as would be the case in almost every situation).	Thank you for your comment. Guidance changed to incorporate ESS into coverage guidance box. Eight trials evaluated the effect of CPAP on neurocognitive or psychological tests, all found significant benefit from CPAP. Reference to impaired cognition has been deleted from the guidance box.
	2	It might be of value to consider whether provider needs to test for alcohol use, as recommendations for abstinence from alcohol is a standard recommendation whether or not a patient is using CPAP.	Evidence source does not address this, except to list avoidance of alcohol as the conservative management arm compared to surgery.
	3	It might also be of value to specify that the provider education should cover avoidance of alcohol, avoidance of CNS-affecting medications, and the contribution of obesity to OSA, when applicable. It could even be required to document (by requesting provider) that a review of medications has been performed, focusing on current use of contraindicated medications, and avoidance of them in the future.	Evidence source does not address this, except to list weight loss, positional therapy, and avoidance of alcohol and sedatives as the conservative management arm compared to surgery. Regarding obesity, three trials of weight loss interventions (primarily diets) found a significant improvement in AHI, ESS and O2 saturation. Regarding provider education, 9 studies evaluated extra support or education to improve compliance with CPAP, however results were inconsistent. Counseling regarding weight loss has been added to the guidance box.
	4	I also believe the literature suggests that compliance with CPAP can be predicted in most cases by usage in the first few weeks, if not sooner. Is there need to have the trial period be 12 weeks-that would seem to be excessive, and given the likely high rate of non-compliance, is a 3 month trial necessary? It seems not, and a significant cost to the system. A shorter trial period might also promote the DME supplier to ensure member awareness of compliance requirements. I would propose a two-stage trial period-the first of 4-6 weeks to establish compliance, and if that first criteria is met, a second criteria at 12-16 weeks to evaluate for effectiveness.	The evidence source identified 5 studies that evaluated predictors of compliance, which included higher AHI, higher ESS score, younger age, snoring, lower CPAP pressure, higher BMI, higher mean oxygen saturation. One of those trials evaluated compliance at 4 weeks and found the only significant predictor to be high baseline AHI. There was a small (3%) decrease in the number of patients compliant with CPAP use between 4 weeks and 12 weeks. No other trials evaluated compliance or predictors of compliance at 4-6 weeks.

Stakeholder	#	Comment	Disposition
	5	It also might be helpful to objectify "effectiveness" or clinical benefit if possible. Thank you for your consideration.	Effectiveness is explained in the text, as follows: "sufficient evidence supporting large improvements in sleep measures with CPAP compared with control (e.g., reducing apnea hypopnea index (AHI), improving symptoms as measured by the Epworth Sleepiness Scale, reducing arousal index, and raising the minimum oxygen saturation). Weak evidence demonstrated no consistent benefit in improving quality of life, neurocognitive measures or other intermediate outcomes."
Industry Location Unknown	6	In response to the draft coverage guidance: Treatment of sleep apnea in adults, I guess my first response would be; is this the full policy? It appears that it may be a summary of medical necessity but does not have guidelines which currently exist in this policy such as when to bill for the sale of the item. For example the current policy has has "a three month trial (rental) period for CPAP is required prior to purchase", the draft does not mention a change in therapy, existing policy states "If a CPAP device was used more than three months and the client is switched to a RAD, then the clinical re-evaluation would occur between the 61st and 91st day following initiation of the RAD".	This document provides general guidance only. Specific implementation of the policy is left to individual payers.
	7	I guess my overall confusion is what is the reasoning for the "draft" is it just in terms of medical appropriateness and nothing further or is the "draft" intended to replace the current rule? If it is intended to replace the current rule it appears to be missing many factors that are vital to providers. Thank you.	Yes, the intent is to address general medical appropriateness, not to replace the current DMAP rule.

HERC Coverage Guidance – Treatment of Sleep Apnea in Adults
Disposition of Public Comments – Round 2

General Comments

Stakeholder	#	Comment	Disposition
CCO Medical Director *Oregon*	1	The proposed Draft Guidance on this matter by HERC is good, as far as it goes. However, some additional specification seems appropriate: While the Guidance specifies the PSG criteria for coverage for a possible 12 week C-pap trial, it does not specify the nature or degree of "clinical benefit" necessary to cover ongoing C-pap use. I propose the following: 1) specify that the post-trial evaluation should include a repeat PSG while on C-pap. As the original criteria for the trial involves this study and it's results, this seems consistent. Otherwise, the "clinical benefit" could be construed to simply be an enrollee's subjective statement of "I feel better". 2) the compliance criteria appear appropriate, but the potential second PSG would help document tolerance of C-pap. 3) the potential second PSG should have specific AHI and RDI criteria for improvement, either specific numbers or % change/improvement. Again this would be consistent with the initial diagnostic criteria.	The WA HTA report does not provide evidence about the clinical significance of changes in AHI. A clinically significant change in ESS is considered to be 1 point (total maximum score = 24). Repeat PSG testing was not addressed in the evidence source. HTAS to consider whether this level of implementation detail is desired in the guidance document. *For HTAS discussion*

HEALTH EVIDENCE REVIEW COMMISSION (HERC)

COVERAGE GUIDANCE: INTENSITY MODULATED RADIATION THERAPY

DRAFT for HTAS Meeting Materials 11/25/2013

HERC COVERAGE GUIDANCE

Intensity modulated radiation therapy (IMRT) is recommended for coverage for these tumor types: prostate cancer (strong recommendation), cervical cancer (weak recommendation).

IMRT is **recommended** for coverage / is **not recommended** for coverage (weak recommendation) for these tumor types and locations: anal cancer, astrocytoma, whole breast radiation, para-aortic lymph node metastases, non-small cell lung cancer, head and neck cancer.

IMRT is not recommended for coverage (weak recommendation) for these malignancy types and locations: esophagus, liver, pancreas, rectum, stomach, brain metastases, glioblastoma, glioma, medulloblastoma, meningioma, pituitary adenoma, partial breast radiation, pleural mesothelioma, small cell lung, sarcoma, sacral chordoma, skin, spinal metastases, thyroid.

IMRT is not recommended for coverage (strong recommendation) for these malignancy types and locations: endometrium and whole pelvis.

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 Heath 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.

DRAFT

EVIDENCE SOURCES

Clark, E.E., Thielke, A., Kriz, H., Bunker, K., Ryan, K., & King, V. (2012). *Intensity modulated radiation therapy.* Olympia, WA: Washington State Health Care Authority Health Technology Assessment Program. Retrieved from http://www.hta.hca.wa.gov/intensity_radiation.html

- *References for included guidelines from Clark 2012 provided in Appendix D*

Glossary Source
U.S. Department of Health and Human Services, National Institutes of Health, & National Cancer Institute. (2010). *Common terminology criteria for adverse events (CTCAE). Version 4.03.* Retrieved October 25, 2013, from http://evs.nci.nih.gov/ftp1/CTCAE/About.html

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

SUMMARY OF EVIDENCE

Clinical Background

Significant advances have been made in the techniques available to deliver external beam radiation therapy (EBRT) as a treatment modality for certain cancers. The goal of these newer techniques is two-fold: to improve the targeting of radiation to the tumor to minimize damage to normal tissue and increase the dose of radiation delivered to the tumor. One of these newer techniques is intensity modulated radiation therapy (IMRT). Intensity modulated radiation therapy, like other forms of EBRT is used as primary treatment or in conjunction with surgery, hormonal therapy and/or chemotherapy; it is used in treatment in primary, recurrent and metastatic cancer.

Intensity modulated radiation therapy has been used for treatment of tumors of the central nervous system, head and neck, breast, prostate, gastrointestinal tract, and gynecologic system. It can be used to treat sites previously treated with radiation and areas in close proximity to organs and vulnerable tissue.

Typically, conventional EBRT [also called 2-dimensional conventional radiation therapy (2DCRT) or 3-dimentional conventional radiation therapy (3DCRT)][1] is delivered in 25 to 50 fractions (doses) delivered five days per week for 5 to 10 weeks. Intensity modulated radiation therapy uses hundreds of radiation beam-shaping devices (collimators) to deliver external beam radiation. The collimators allow the intensity of the radiation to

[1] In this report 2DCRT and 3DCRT are grouped together as conventional radiation therapy (CRT) except where individual studies compare IMRT to either 2DCRT or 3DCRT. Current conventional EBRT is also referred to as CRT.

vary during a treatment session thus different doses of radiation can be directed at different areas of the tumor and nearby tissues. The goal of IMRT is to increase radiation to the tumor while reducing radiation exposure to normal tissue. Treatment planning for IMRT involves additional steps to contour the radiation dose to the tumor and to reduce radiation to surrounding tissues; these additional steps and the additional time required to administer each treatment fraction during the course of therapy result in considerably more cost for IMRT than EBRT. Intensity modulated radiation therapy typically costs approximately twice as much as EBRT. The following approximate charges for IMRT were identified in the source report:

- Non-small cell lung cancer: EBRT$55,000, IMRT $146,000
- Breast cancer (whole breast irradiation): EBRT $9,500, IMRT $17,900
- Breast cancer (partial breast irradiation): EBRT $7,200, IMRT $9,200
- Prostate cancer: EBRT ranging from $10,000 to $27,000, IMRT ranging from $33,000 to $52,000

Evidence Review

The source report presents evidence about the use of IMRT for malignancies in the following anatomic locations: abdomen (anal/rectal, liver, and pancreas), brain, breast, female pelvis, head and neck, lung, prostate, soft tissue sarcomas, and other cancer sites (skin, thyroid, spinal metastases). Sixteen systematic reviews and 108 individual studies met inclusion criteria. The majority of studies were non-comparative and in adults. Only two studies for medulloblastoma were exclusively in the pediatric population. Overall, there is limited evidence to answer many of the key questions and the populations were heterogeneous.

Efficacy

The overall strength of evidence for efficacy outcomes [*e.g.,* overall survival (OS), disease-specific survival (DSS), disease-free survival (DFS), progression-free survival (PFS), recurrence, quality of life (QoL), harms, etc.] ranged from moderate to very low with most being low to very low. In general, for patient survival and recurrence outcomes, the results were heterogeneous, and for many locations, there was no comparative data. Therefore, no general conclusions can be drawn for patient survival and recurrence outcomes.

The findings for QoL were inconsistent except in two anatomic locations with moderate overall strength of evidence findings. The first is whole breast irradiation, in which there were no differences in QoL for IMRT compared to standard radiation therapy (EBRT). The second is head and neck cancers, which found an improvement in overall QoL for IMRT compared to 2D- and 3DCRT.

Harms

Harms were mostly regional toxicities based on the location of the malignancy and commonly included acute and late toxicities [*e.g.*, gastrointestinal (GI), genitourinary (GU), xerostomia, skin, pneumonitis, esophagitis, etc.][2]. There was moderate strength of evidence findings for two outcomes for whole breast irradiation, one outcome for head and neck cancer and one outcome for prostate cancer. For whole breast irradiation, there was moderate strength of evidence that the EBRT group was more likely to develop any Grade of telangiectasia compared to patients who received IMRT. In addition, there was moderate strength of evidence that there were no significant differences in acute toxicities (Grade 2 or higher, Grade 3 or 4 skin toxicities) for IMRT compared to EBRT for whole breast irradiation. For head and neck cancer, there was moderate strength of evidence that IMRT reduces Grade 2 or greater xerostomia compared to EBRT. For prostate cancer, there was a moderate strength of evidence that IMRT improves gastrointestinal toxicities compared to EBRT. Deaths and serious adverse events (*e.g.,* harms requiring surgery) were not common, but were reported by a few studies across several anatomic cancer locations.

Details of this information by tumor site are presented in the tables that follow. For any cancer type that has more than a very low strength of evidence, additional detail is provided.

Abdomen (Anus, Esophagus, Liver, Pancreas, Rectum, Stomach, Whole Pelvis)

Cancer Type	Strength of Evidence	Efficacy Outcomes	Harms Outcomes
Anal	Very low (2 case series, 1 cohort study, N=97)	Better 3 year OS, PFS with IMRT	Reduced acute nonhematologic toxicity in genital area and diarrhea with IMRT
Esophagus	Very low (1 case series, n=30)	Unknown	Unknown
Liver	Very low (3 case series, N=70)	Unknown	Unknown
Pancreas	Very low (1 case series, n=47)	Unknown	Unknown
Rectum	Very low (1 case	Unknown	Unknown

[2] Toxicities are graded as follows: *Grade 1:* Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated. *Grade 2:* Moderate; minimal, local or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living (ADL) (*e.g.*, preparing meals, shopping for groceries or clothes, using the telephone, managing money, etc.). *Grade 3:* Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care ADL (*e.g.,* bathing, dressing and undressing, feeding self, using the toilet, taking medications, and not bedridden). *Grade 4:* Life-threatening consequences; urgent intervention indicated. *Grade 5:* Death related to adverse events.

Cancer Type	Strength of Evidence	Efficacy Outcomes	Harms Outcomes
	series, n=63)		
Stomach	Very low (2 cohort studies, N=111)	Inconsistent	Less impact on renal function with IMRT
Whole pelvis	No studies on effectiveness, 1 cohort study for harms (n=134)	Unknown	No difference in toxicity

N = total number of participants in all studies; n = number of participants in one study

Pancreatic Cancer

One cost study was identified for pancreatic cancer that compared gemcitabine in combination with stereotactic body radiation therapy, IMRT or EBRT. The authors calculated an incremental cost effectiveness ratio (ICER) of $1,584,100 for gemcitabine plus IMRT compared to gemcitabine plus EBRT.

Brain (Astrocytoma, Brain Metastases, Glioblastoma Multiforme, High-grade Glioma, Medulloblastoma, Meningioma, Pituitary Adenoma)

Cancer Type	Strength of Evidence	Efficacy Outcomes	Harms Outcomes
Astrocytoma	Very low (1 cohort study, n=85)	Better 1 and 2 year OS, PFS with IMRT	Unknown
Brain mets	Very low (1 case series, n=29)	Unknown	Unknown
Glioblastoma multiforme	Very low (8 case series, N not stated, largest n=42)	Unknown	Unknown
High-grade glioma	Very low (2 case series, N=98)	Unknown	Unknown
Medulloblastoma	Very low (1 case series, n=33)	Unknown	Reduced Grade 3 and 4 ototoxicity with IMRT, no diff in cognitive function
Meningioma	Very low (3 case series, N=257)	Unknown	Unknown
Pituitary adenoma	Very low (1 case series, n=34)	Unknown	Unknown

Breast (Whole Breast Radiation, Partial Breast Radiation)

Cancer Type	Strength of Evidence	Efficacy Outcomes	Harms Outcomes
Whole breast radiation	Low (19 total studies, N>1,000)	Inconsistent findings for OS, DSS; no difference in recurrence	Inconsistent findings for cosmesis
Whole breast radiation	Moderate (see above)	No sig difference in QoL	No sig difference in acute toxicity, telangiectasia less likely with IMRT
Partial breast radiation	Very low (3 case series, N=175)	Unknown	Unknown

Breast Cancer: Whole Breast Radiation

For whole breast radiation of breast cancer, two systematic reviews reported inconclusive findings for patient survival and one retrospective cohort study (N=240) comparing IMRT to 2DCRT reported no significant differences in overall survival. The overall strength of evidence is low that there are inconsistent findings for patient survival (overall survival, disease-specific survival).

For cancer recurrence and distant metastases, the only comparative study reported no significant differences compared to 2DCRT. All other studies were non-comparative, and reported a range of 0% to 2.9% for local recurrence over three years. The one study that reported on distant metastasis reported a rate of 2.7%, but the time period was not specified. The overall strength of the evidence for these outcomes is low.

There is limited evidence on QoL outcomes from IMRT. There is moderate overall strength of evidence that IMRT compared to EBRT does not result in significant differences in QoL.

The overall strength of evidence is low that there is inconsistent evidence for breast cosmesis when IMRT is compared to EBRT. There is moderate overall strength of evidence that IMRT compared to EBRT does not result in a significant difference in acute toxicities (*i.e.,* Grade 2 or higher acute toxicities, Grade 3 or 4 skin toxicities).

One large prospective randomized controlled trial (n=815) reported that the EBRT group was 1.68 times more likely to develop any Grade (1, 2, or 3) of telangiectasia compared to IMRT (moderate overall strength of evidence). There are inconsistent findings that IMRT, compared to EBRT, is associated with lower rates of late Grade 2 or greater breast edema or hyperpigmentation (low overall strength of evidence). Limited evidence reported no significant differences in late Grade 2 or greater fat necrosis or induration/fibrosis for IMRT compared to EBRT; the overall strength of evidence is low.

Two cost studies were identified. One, published in 2011, analyzed the Surveillance, Epidemiology and End Results (SEER) - Medicare database and found the average cost of EBRT was $7,179 while the average cost of IMRT was $15,230. The second modeled several treatment regimens for stage one breast cancer. Total direct medical costs for EBRT ranged from $6,100 to $10,900, while total mean costs for IMRT were $19,300.

Female Pelvis (Cervical, Endometrial, Para-aortic Lymph Node Metastases)

Cancer Type	Strength of Evidence	Efficacy Outcomes	Harms Outcomes
Cervical	Low (3 cohort studies, 3 case series, N=928)	Better OS, PFS with IMRT	Lower frequency of toxicities with IMRT
Endometrial	Very low (1 cohort study, n=33)	Unknown	No difference in toxicities
Para-aortic lymph node mets	Very low (1 cohort study, 1 case series, N=89)	Better 2 and 3 year OS with IMRT	Less frequent GI and GU toxicities with IMRT

Cervical Cancer

For treatment of cervical cancer with IMRT, overall survival findings are inconsistent. Although two smaller cohort and case series studies found no difference compared to EBRT, one larger cohort study included in a systematic review found significant benefit for patients treated by IMRT. The overall strength of evidence is low that IMRT was associated with increased disease-specific survival and overall survival for patients with cervical cancer compared to EBRT.

With regard to harms, findings from one cohort and two case series studies provide an overall low strength of evidence that IMRT was associated with lower frequency of late gastrointestinal and genitourinary toxicities (cystitis, proctitis, enterocolitis) than EBRT.

Lung (Non-small Cell, Pleural Mesothelioma, Small Cell)

Cancer Type	Strength of Evidence	Efficacy Outcomes	Harms outcomes
Non-small cell	Low (1 cohort study, 5 case series, N=785)	Better OS, no difference in locoregional PFS or distant mets-free survival with IMRT	Lower levels of Grade 3 or worse pneumonitis with IMRT
Pleural mesothelioma	Very low (3 case series, N= not stated)	Better local recurrence rate	Unknown

Cancer Type	Strength of Evidence	Efficacy Outcomes	Harms outcomes
Small cell	Very low (1 case series, n=60)	Unknown	Unknown

Non-Small Cell Lung Cancer

For non-small cell lung cancer (NSCLC), one comparative study and five case series were identified. The overall strength of the evidence is low that patients treated with IMRT compared to 3DCRT for non-small cell lung cancer had better overall survival. The overall strength of evidence is low that there were no significant differences in distant metastasis-free survival or locoregional PFS for IMRT compared to 3DCRT. No conclusions can be drawn from the case series since they did not compare IMRT to EBRT. The overall strength of evidence for all other outcomes is very low.

With regard to harms, two comparative studies and six case series were identified. The overall strength of the evidence is low that NSCLC patients treated with IMRT compared to EBRT had significantly lower levels of greater than or equal to Grade 3 pneumonitis.

The remaining outcomes were only reported in noncomparative studies, and therefore no conclusions can be drawn for IMRT compared to other treatments. In general, Grade 1 or 2 toxicities were reported with varying percentages. The good to fair quality case series reported patients with esophagitis, Grade 2 and 3 late esophageal stricture, Grade 3 or greater treatment-related pneumonitis[3] (including one death), pulmonary fibrosis, and Grade 3 pulmonary toxicities. The overall strength of evidence is very low.

Other (Head and Neck, Prostate, Sarcoma, Sacral Chordoma, Skin, Thyroid, Spinal Metastases)

Cancer Type	Strength of Evidence	Efficacy Outcomes	Harms Outcomes
Head and neck	Very Low (1 SR, 1 cohort study, N not stated)		No difference in osteonecrosis or hearing loss
Head and neck	Low (multiple studies including at least 3 RCTs, at least 14 cohort studies, N not stated	No difference in local tumor control or OS	
Head and neck	Moderate (see	Better xerostomia-related	Less Grade 2 or worse

[3] Grade 3 pneumonitis is defined as inflammation focally or diffusely affecting lung parenchyma and causing severe symptoms, limiting self-care for ADLs, with oxygen indicated.

Cancer Type	Strength of Evidence	Efficacy Outcomes	Harms Outcomes
	above)	QoL with IMRT	xerostomia with IMRT
Prostate	Low (8 cohort studies, N not stated)	No difference in OS, better biochemical DFS (5 years) and recurrence (3 years)	Fewer GU toxicities, hip fracture less likely with IMRT, erectile dysfunction inconsistent
Prostate	Moderate (at least 15 cohort studies, N not stated)		Fewer GI toxicities with IMRT
Sarcoma	Very low (1 case series, n=27)	Unknown	Unknown
Sacral chordoma	Very low (1 case series, n=34)	Unknown	Unknown
Skin	Very low (1 case series, n=21)	Unknown	Unknown
Spinal mets	Very low (3 case series, N=220)	Unknown	Unknown
Thyroid	Low (1 cohort study, 1 case series, N=151)	No difference in all survival measures, fewer late morbidities with IMRT	Unknown

Head and Neck Cancer

For head and neck cancer, the overall strength of evidence is low that there was no significant difference between IMRT and EBRT in local tumor control or overall survival. The findings on xerostomia-related QoL are inconsistent. However, there is a preponderance of the evidence supporting that IMRT compared to 2D- and 3DCRT improves xerostomia-related QoL. Therefore, the overall strength of evidence that IMRT compared to 2D- and 3DCRT improves xerostomia-related QoL is moderate.

Six systematic reviews and an additional 49 articles address harms. There is moderate overall strength of evidence that IMRT reduces Grade 2 or greater xerostomia compared to EBRT. There is a very low strength of evidence that there is no significant difference in incidence of osteonecrosis from IMRT compared to EBRT. There is very low strength of evidence that there is no significant difference in hearing loss from IMRT compared to EBRT. The overall strength of evidence for all other harms (*i.e.,* nausea, vomiting, fatigue, dermatitis, mucositis, dysphagia, laryngeal symptoms) is very low.

One cost study was identified, which reported costs associated with IMRT at nine different centers in France. Authors report that costs ranged from €6,332 at experienced centers to €14,192 at new centers. No comparison was made to EBRT.

Prostate Cancer

For prostate cancer, three systematic reviews and seven cohort studies were identified. There is low strength of evidence that there were no significant differences in overall survival for IMRT compared to EBRT at 30 months. There was low overall strength of evidence for a significant difference in biochemical disease-free survival at 60 months favoring the IMRT group compared to EBRT. There is low strength of evidence that IMRT compared to EBRT had lower rates of cancer recurrence at three years.

Two fair quality cohort studies reported inconsistent findings for QoL in different populations. Therefore, no conclusions can be drawn, and the overall strength of evidence is low.

Comparison of harms is difficult because of the different dosages, treatment regimens, cancer stages, and outcomes studied. However, based on three large cohorts, there is an overall moderate strength of evidence that IMRT improves gastrointestinal toxicities compared to EBRT. There is an overall low strength of evidence that IMRT improves genitourinary toxicities compared to EBRT.

There is low strength of evidence that the IMRT group was less likely to experience hip fractures compared to conventional radiation therapy. Based on four cohort studies, the evidence on erectile dysfunction is inconsistent. A large, good quality cohort study found that the IMRT group was more likely to receive a diagnosis of erectile dysfunction (RR 1.12; 95% CI 1.03-1.20). However, the effect size was small. There is an overall low strength of evidence for this outcome.

One technology assessment reported on cost effectiveness. Three different cost-effectiveness analyses were included. Two of the three assumed improved survival with IMRT, and resulted in ICERs of $16,182 and $42,450. The third analysis did not assume improved survival and calculated an ICER of $706,000. Since the evidence does not support improved survival, the third analysis is likely more accurate. A second study evaluated the two-year US national cost of various initial treatment strategies for prostate cancer (active surveillance, radical prostatectomy, brachytherapy, EBRT and IMRT). Adjusted two-year costs ranged from $21,400 for active surveillance to $68,300 for IMRT.

Thyroid Cancer

For thyroid cancer, the overall strength of evidence is low that there were no significant differences in all survival measures for IMRT compared to EBRT. There is low overall strength of evidence that IMRT had less late morbidities than the EBRT group. There are few comparative studies addressing other harms and therefore no conclusions can be reached comparing IMRT to other treatments.

DRAFT

Subgroups and Costs

There was insufficient evidence to address differential safety and efficacy for any subgroup. Cost studies were identified for pancreatic cancer, breast cancer, head and neck cancer and prostate cancer, and are discussed in those sections. For all other malignancy locations, no cost information was identified and the evidence is insufficient.

Guidelines

The WA HTA report identified a total of 17 guidelines and 11 appropriateness criteria (the latter from the American College of Radiology) that addressed IMRT. Those pertaining to specific cancers are summarized below. All of the National Comprehensive Cancer Network (NCCN) guidelines were rated by report authors as poor quality. They state:

> "While the NCCN guidelines have a transparent guideline development process and are explicit about guideline panel members and NCCN staff conflicts of interest, the methods for identifying and selecting evidence are unclear. After several email conversations with NCCN staff about their methodology, it is still unclear how evidence is identified (e.g., search strategy and databases searched), what the inclusion/exclusion criteria are, and if individual studies are assessed for quality. Based on the dearth of information in these areas, all of the NCCN guidelines were rated as poor."

The ACR Appropriateness Criteria® are developed through an expert panel process and focus on diagnostic imaging, interventional radiology, and radiation oncology. Technologies are given an appropriateness rating between 1 and 9; the appropriateness rating can vary depending on treatment situation and patient characteristics. Ratings of 1, 2 or 3 are considered usually not appropriate, ratings of 4, 5 or 6 are considered as may be appropriate, and ratings of 7, 8, or 9 are considered usually appropriate. All of the ACR Appropriateness Criteria® included in the WA HTA report were rated by report authors as fair quality.

<u>Anal Cancer</u>: The NCCN states that IMRT may be used in place of 3D conformal radiation therapy, and that IMRT requires expertise and careful target design. The ACR gives IMRT an Appropriateness Criteria® of 6 and "cautiously recommends" the use of IMRT if performed outside of a study protocol setting.

<u>Breast Cancer</u>: The NCCN provides recommendations about IMRT in their local-regional treatment guidelines for stage I, IIA, IIB, or T2N1M0 invasive breast cancer. Tissue wedging, forward planning with segments (step and shoot), or IMRT is recommended.

Central Nervous System Cancers: The NCCN states that 3D-planning or IMRT may be appropriate for low-grade astrocytomas and oligodendrogliomas.

Cervical Cancer: The NCCN suggests that IMRT may be helpful, but that it is not to be used as a routine alternative to brachytherapy. The guideline stresses that very close attention to detail and reproducibility is needed. The ACR gives IMRT an Appropriateness Criteria® score of 3 to 8 depending on case variability. However, the ACR states that IMRT is "not indicated for routine treatment of cervical cancer". In addition, the ACR gives an Appropriateness Criteria® score of 7 to IMRT as adjuvant therapy in the management of early stage cervical cancer and notes that great care is required in delineation of clinical target volume.

Colon Cancer: The NCCN suggests that IMRT be reserved for unique clinical situations including re-irradiation of previously treated patients with recurrent disease.

Esophageal and Esophagogastric Junction Cancers: The NCCN states that IMRT may be appropriate in selected cases to reduce dose to normal structures.

Gastric Cancer: The NCCN states that IMRT may be appropriate in selected cases to reduce dose to normal structures.

Head and Neck Cancers: The NCCN states that either 3D conformal radiation therapy or IMRT may be appropriate in the treatment of head and neck cancers, and that IMRT may be used at the discretion of the treating physician. Specifically for resectable oropharyngeal squamous cell carcinoma, the ACR gives IMRT an Appropriateness Criteria® score of 8 to 9 depending on case variability.

Lung Cancer: Two guidelines for lung cancer were identified: one on malignant pleural mesothelioma and one on NSCLC, with an additional three ACR Appropriateness Criteria® documents identified for NSCLC. For mesothelioma, the NCCN states that "IMRT […] should only be used in experienced centers or on protocol" and when IMRT is applied, that the NCI/ASTRO IMRT guidelines should be strictly followed. The NCCN guideline further states that "special attention should be paid to minimize radiation to the contralateral lung" and that the "mean lung dose should be kept as low as possible".

For NSCLC, the NCCN states that CT-planned 3DCRT is the minimum standard for treatment and that "use of more advanced technologies [including IMRT] is appropriate when needed to deliver adequate tumor doses while respecting normal tissue dose constraints". The ACR has varying Appropriateness Criteria® for the use of IMRT for NSCLC. The ACR Appropriateness Criteria® for IMRT for postoperative adjuvant therapy for NSCLC is a 6, and is an 8 (with tumor motion strategy required in addition to strict dosimetric criteria) for nonsurgical treatment and induction and adjuvant therapy for N2 for NSCLC.

For SCLC, the NCCN states that the use of IMRT may be considered.

Non-spine Bone Metastases: The ACR gives IMRT an Appropriateness Criteria® score of two for non-spine bone metastases.

Prostate Cancer: The NCCN states that 3D conformal and IMRT techniques should be employed if radiation treatment is being considered for prostate cancer. For T1 and T2 prostate cancer, the ACR gives IMRT an Appropriateness Criteria® score of 8; for postradical prostatectomy irradiation in prostate cancer, the ACR gives IMRT an Appropriateness Criteria® score of 2 to 8, depending on case variability.

Rectal Cancer: The NCCN states that IMRT should only be used in clinical trial settings or in unique clinical situations for rectal cancer. The ACR gives IMRT an Appropriateness Criteria® score of 1 for rectal cancer and states that it should be for investigational use only.

Testicular Cancer: The NCCN does not recommend IMRT for the treatment of testicular cancer.

Thyomas: The NCCN states that IMRT may "further improve the dose distribution and decrease dose to the normal tissue as indicated" and that if IMRT is applied, the NCI/ASTRO guidelines should be strictly followed.

Summary of NCCN and ACR Guidelines by Malignancy

Malignancy	Guideline (Year) Quality	Usually Not Appropriate / Not Recommended	May be Appropriate	Usually Appropriate / Recommended
Abdomen				
Anal/Rectal carcinoma	NCCN (2012a) Poor		IMRT may be used in place of 3D conformal RT. Requires expertise and careful target design.	
Anal cancer	Poggi [ACR] (2010) Fair		ACR 6 "cautiously recommends" the use of IMRT if performed outside of a protocol setting.	
Colon cancer	NCCN (2012e) Poor	IMRT reserved only for unique clinical situations including re-irradiation of previously treated patients with recurrent disease.		
Esophageal and esophagogastric junction cancers	NCCN (2012f) Poor		In selected cases to reduce dose to normal structures.	

Malignancy	Guideline (Year) Quality	Usually Not Appropriate / Not Recommended	May be Appropriate	Usually Appropriate / Recommended
Gastric cancers	NCCN (2012g) Poor		In selected cases to reduce dose to normal structures.	
Rectal cancer	NCCN (2012m) Poor	IMRT should only be used in clinical trial setting or in unique clinical situations including re-irradiation of recurrent disease after previous radiotherapy).		
Resectable rectal cancer	Suh [ACR] (2007) Fair	ACR 1 (investigational use only)		
Brain				
Central nervous system	NCCN (2012c) Poor		For low-grade astrocytoma and oligodendroglioma. 3D planning or IMRT.	
Breast				
Breast cancer	NCCN (2012b) Poor			Recommended following CT-based treatment planning
Female Pelvis				
Cervical cancer	Gaffney [ACR] (2010) Fair	"not indicated for routine treatment of cervical cancer" ACR (3-8)		
Cervical cancer	NCCN (2012d) Poor		May be helpful. Not to be used as routine alternative to brachytherapy. Very close attention to detail and reproducibility needed.	
Cervical cancer (role of adjuvant therapy in the management of early stage)	Wolfson [ACR] (2011) Fair			ACR 7 (great care required in delineation of CTV)
Head and Neck				
Head and neck cancers	NCCN (2012h) Poor		Either 3D conformal RT or IMRT. IMRT may be used at the discretion of treating physicians.	
Head and neck	NCCN		IMRT, 3D and 2D conformal	

DRAFT

Malignancy	Guideline (Year) Quality	Usually Not Appropriate / Not Recommended	May be Appropriate	Usually Appropriate / Recommended
cancers (mucosal melanoma)	(2012j) Poor		techniques may be used as appropriate. IMRT may be used at the discretion of treating physicians.	
Resectable oropharyngeal squamous cell carcinoma	Quon [ACR] (2010) Fair			Dependent on patient characteristics, ACR 8-9
Lung				
Malignant pleural mesothelioma	NCCN (2012i) Poor		IMRT should only be used in experienced centers or on protocol. NCI/ASTRO IMRT guidelines should be strictly followed.	
NSCLC (postoperative adjuvant therapy)	Decker [ACR] (2011) Fair		Dependent on patient characteristics and tumor stage, ACR 6	
NSCLC (nonsurgical treatment)	Gewanter [ACR] (2010) Fair			ACR 8 (with tumor motion strategy required in addition to strict dosimetric criteria)
NSCLC (induction and adjuvant therapy for N2)	Gopal [ACR] (2010) Fair			ACR 8 (with tumor motion strategy required in addition to strict dosimetric criteria)
NSCLC	NCCN (2012k) Poor		Use of IMRT appropriate when need to deliver adequate tumor doses while respecting normal tissue dose constraints.	
SCLC	NCCN (2012n) Poor		In selected pts, IMRT may be considered.	
Prostate				
Prostate cancer (T1 and T2)	Morgan [ACR] (2011) Fair			Dependent on patient characteristics, ACR 8

DRAFT

Malignancy	Guideline (Year) Quality	Usually Not Appropriate / Not Recommended	May be Appropriate	Usually Appropriate / Recommended
Prostate cancer	NCCN (2012l) Poor			3D conformal or IMRT (no preference given)
Postradical prostatectomy irradiation in prostate cancer	Rossi [ACR] (2010) Fair	Dependent on patient characteristics, ACR 2-8		Dependent on patient characteristics, ACR 2-8
Other Cancers				
Non-spine bone metastases	Lutz [ACR] (2011) Fair	Dependent on patient characteristics, ACR 2		
Testicular cancer	NCCN (2012o) Poor	IMRT not recommended		
Thymomas and thymic carcinomas	NCCN (2012p) Poor		IMRT may further improve dose distribution and decrease dose to the normal tissues as indicated. Strictly follow NCI/ASTRO IMRT guidelines.	

Summary of Guidelines

The NCCN guidelines and the ACR Appropriateness Criteria® are consistent in their statements and recommendations for IMRT for anal, prostate, and rectal cancer. There are no ACR Appropriateness Criteria® ratings for breast, central nervous system, colon, esophageal and esophagogastric junction, gastric, general head and neck, mesothelioma, testicular and thymic cancers. There are no NCCN guidelines for non-spinal bone metastases. Based on poor to fair quality guidelines, IMRT is considered usually appropriate by the ACR and/or recommended by the NCCN for breast cancer, resectable oropharngeal squamous cell carcinoma, nonsurgical treatment of NSCLC, induction and adjuvant therapy for N2 NSCLC, and prostate cancer. Intensity modulated radiation therapy is not recommended by the NCCN or considered appropriate by the ACR for the treatment of colon cancer, rectal cancer, non-spine bone metastases, and testicular cancer. For cervical cancer, the NCCN and the ACR have inconsistent recommendations ranging from usually not appropriate/not recommended to usually appropriate/recommended. For all other cancers discussed, IMRT is considered as a possible appropriate form of treatment by the ACR and NCCN.

[Evidence Source]

DRAFT

Evidence Summary

In general, for patient survival and recurrence outcomes, the results were heterogeneous, and for many locations, there was no comparative data. Therefore, no general conclusions can be drawn for patient survival and recurrence outcomes. The findings for QoL were inconsistent except in two anatomic locations. The first is whole breast irradiation, in which there is moderate strength of evidence that there are no differences in QoL for IMRT compared to standard radiation therapy (EBRT). The second is head and neck cancers, which found an improvement in overall QoL for IMRT compared to 2D- and 3DCRT. Harms were mostly regional toxicities based on the location of the malignancy. There was moderate strength of evidence that for whole breast irradiation, EBRT is more likely to result in any Grade of telangiectasia compared to IMRT, and there are no significant differences in acute toxicities. For head and neck cancer, there was moderate strength of evidence that IMRT reduces Grade 2 or greater xerostomia compared to EBRT. For prostate cancer, there was a moderate strength of evidence that IMRT improves gastrointestinal toxicities compared to EBRT. Cost studies were identified for pancreas, breast, head and neck, and prostate cancers, and consistently reported that IMRT costs more than other treatments (approximately twice as much). For those cancers in which an ICER was calculated (pancreas, prostate), IMRT was not found to be cost effective.

DRAFT

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	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Values and preferences	Coverage Recommendation
IMRT for prostate cancer	Fewer GI toxicities compared to EBRT	Moderate	Moderate to high	Low variability	Recommended for coverage *(strong recommendation)*
IMRT for cervical cancer	Better OS, PFS compared to EBRT	Low	Moderate to high	Low variability	Recommended for coverage *(weak recommendation)*
Whole Breast IMRT for breast cancer	No difference in effectiveness or harms (except telangiectasia) compared to EBRT	Moderate	Moderate to high	Low variability	To be determined
IMRT for anal cancer, astrocytoma, para-aortic lymph node metastases	No clear differences in overall balance compared to EBRT	Very low	Moderate to high	Low variability	To be determined
IMRT for non-small cell lung cancer	No clear differences in overall balance compared to EBRT	Low	Moderate to high	Low variability	To be determined
IMRT for head and neck cancer	Better QoL, less xerostomia compared to EBRT	Moderate	Moderate to high	Low variability	To be determined
	No difference in local tumor control or OS	Low	Moderate to high	Low variability	
	No difference in OS, improved bDFS[4] at 60 months	Low	Moderate to high	Low variability	
Partial Breast IMRT for breast cancer	Unknown	Very Low	Moderate	Low variability	Not recommended for coverage (weak recommendation)
IMRT for all other cancers	Unknown	Very low	Moderate to high	Low variability	Not recommended for coverage (weak recommendation)

*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

[4] bDFS = biochemical disease-free survival

DRAFT

POLICY LANDSCAPE

No quality measures pertaining to IMRT were identified when searching the National Quality Measures Clearinghouse. With regard to payer coverage policies, the WA HTA report identified the following:

There are no Medicare national coverage decisions (NCDs) for IMRT. Coverage policies are left to regional Medicare contractors. The regional contractor for Washington State has issued two localized coverage determinations (LCDs) for IMRT (L24318 and L31415). Both LCDs cover brain tumors, brain metastasis, prostate cancer, lung cancer, pancreas cancer and other upper abdominal sites, spinal cord tumors, head and neck cancer, adrenal tumors, and pituitary tumors. An additional LCD (L30316) that covers 40 states (including Washington) states IMRT is indicated as a standard treatment option for central nervous system tumors (including brain and spinal cord), head and neck cancers, prostate cancer, selected cases of thoracic and abdominal malignancies, selected cases of breast cancers (with close proximity to critical structures), and pelvic and retroperitoneal tumors. Clinical criteria for medical necessity of IMRT include adjacent critical structures that need to be protected; areas adjacent to a previously irradiated area; concave or convex target volume; and radiation doses in excess of those utilized with conventional treatment.

Aetna
For approval of IMRT, Aetna requires that critical structures located close to tumors cannot be adequately protected using conventional EBRT. Specific tumor types are not listed.

GroupHealth
GroupHealth does not require a medical necessity review for use of IMRT in head and neck and prostate cancers.

Regence BCBS
Four unique policies are used by Regence BCBS for coverage determinations. Treatment may be considered medically necessary for the treatment of squamous cell cancer of the anal canal and head and neck cancers. In primary prostate cancer IMRT can be used as the main treatment, as a salvage treatment for failed main treatment or recurrence and as adjuvant therapy immediately following prostatectomy. The use of IMRT for the treatment of breast, lung, and other abdominal or pelvic tumors is considered medically necessary when there is prior radiation to the area, or there are critical structures in the radiation field. For other cancers, IMRT is not considered medically necessary.

DRAFT

COMMITTEE DELIBERATIONS – HTAS

COMMITTEE DELIBERATIONS – VBBS

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/outcome[5]

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.

[5] Includes risk of bias, precision, directness, consistency and publication bias

DRAFT

Appendix B. Applicable Codes

CODES	DESCRIPTION
ICD-9 Diagnosis Codes	
140	Malignant neoplasm of lip
141	Malignant neoplasm of tongue
142	Malignant neoplasm of major salivary glands
143	Malignant neoplasm of gum
144	Malignant neoplasm of floor of mouth
145	Malignant neoplasm of other and unspecified parts of mouth
146	Malignant neoplasm of oropharynx
147	Malignant neoplasm of nasopharynx
148	Malignant neoplasm of hypopharynx
149	Malignant neoplasm of other and ill-defined sites within the lip oral cavity and pharynx
150	Malignant neoplasm of esophagus
151	Malignant neoplasm of stomach
154	Malignant neoplasm of rectum and anus
155.0	Malignant neoplasm of liver, primary
155.2	Malignant neoplasm of liver, not specified as primary or secondary
157	Malignant neoplasm of pancreas
160	Malignant neoplasm of nasal cavities middle ear and accessory sinuses
161	Malignant neoplasm of larynx
162	Malignant neoplasm of trachea bronchus and lung
163	Malignant neoplasm of pleura
170.2	Malignant neoplasm of bone, vertebral column excluding sacrum and coccyx
170.6	Malignant neoplasm of bone, pelvic bones, sacrum and coccyx
171	Malignant neoplasm of connective and other soft tissue
172	Malignant melanoma of skin
173	Other malignant neoplasm of skin
174	Malignant neoplasm of breast
180	Malignant neoplasm of cervix uteri
182	Malignant neoplasm of body of uterus
184	Malignant neoplasm of other/unspecified female genital organs
185	Malignant neoplasm of prostate
190	Malignant neoplasm of eye
191	Malignant neoplasm of brain
192	Malignant neoplasm of other and unspecified parts of nervous system
193	Malignant neoplasm of thyroid gland
196	Secondary and unspecified malignant neoplasm of lymph nodes
197.5	Secondary malignant neoplasm of large intestine and rectum
197.7	Malignant neoplasm of liver, secondary
197.8	Secondary malignant neoplasm of other digestive organs and spleen
198.3	Secondary malignant neoplasm of brain and spinal cord
198.7	Secondary malignant neoplasm of adrenal gland
198.89	Secondary malignant neoplasm of other specified sites
225	Benign neoplasm of brain and other parts of nervous system
227.3	Benign neoplasm of pituitary gland and craniopharyngeal duct

DRAFT

CODES	DESCRIPTION
ICD-10 Diagnosis Codes	
C00-14	Malignant neoplasms of lip, oral cavity and pharynx
C15	Malignant neoplasm of esophagus
C16	Malignant neoplasm of stomach
C18	Malignant neoplasm of colon
C19	Malignant neoplasm of rectosigmoid junction
C20	Malignant neoplasm of rectum
C21	Malignant neoplasm of anus and anal canal
C22	Malignant neoplasm of liver and intrahepatic bile ducts
C25	Malignant neoplasm of pancreas
C30	Malignant neoplasm of nasal cavity and middle ear
C31	Malignant neoplasm of accessory sinuses
C32	Malignant neoplasm of larynx
C33	Malignant neoplasm of trachea
C34	Malignant neoplasm of bronchus and lung
C41.2	Malignant neoplasm of bone, vertebral column excluding sacrum and coccyx
C41.4	Malignant neoplasm of bone, pelvic bones, sacrum and coccyx
C43	Malignant melanoma of skin
C44	Other malignant neoplasm of skin
C45	Mesothelioma
C49	Malignant neoplasm of connective and other soft tissue
C50	Malignant neoplasm of breast
C53	Malignant neoplasm of cervix uteri
C54	Malignant neoplasm of body of uterus
C57	Malignant neoplasm of other/unspecified female genital organs
C61	Malignant neoplasm of prostate
C69-72	Malignant neoplasms of eye, brain and other parts of central nervous system
C73	Malignant neoplasm of thyroid gland
C74	Malignant neoplasm of adrenal gland
C76	Malignant neoplasm of other and ill-defined sites
C79	Secondary malignant neoplasm of other and unspecified sites
D33	Benign neoplasm of brain and other parts of central nervous system
D35	Benign neoplasm of other and unspecified endocrine glands
ICD-9 Volume 3 (Procedure Codes)	
92.29	Other radiotherapeutic procedure
CPT Codes	
77338	Multi-leaf collimator (MLC) device(s) built for IMRT per IMRT plan
77418	Intensity modulated treatment delivery, single or multiple fields, via narrow spatially and temporally modulated beams, per treatment session
77301	Intensity modulated radiotherapy plan, including dose-volume histograms for target and critical structure partial tolerance specifications
77014	Computed tomography guidance for placement of radiation therapy fields
77261/2/3	Radiation Therapy Planning, simple, intermediate, complex
77280/85 77290/95	Set radiation therapy field, simple, intermediate, complex (0) or 3 dimensional (5)

DRAFT

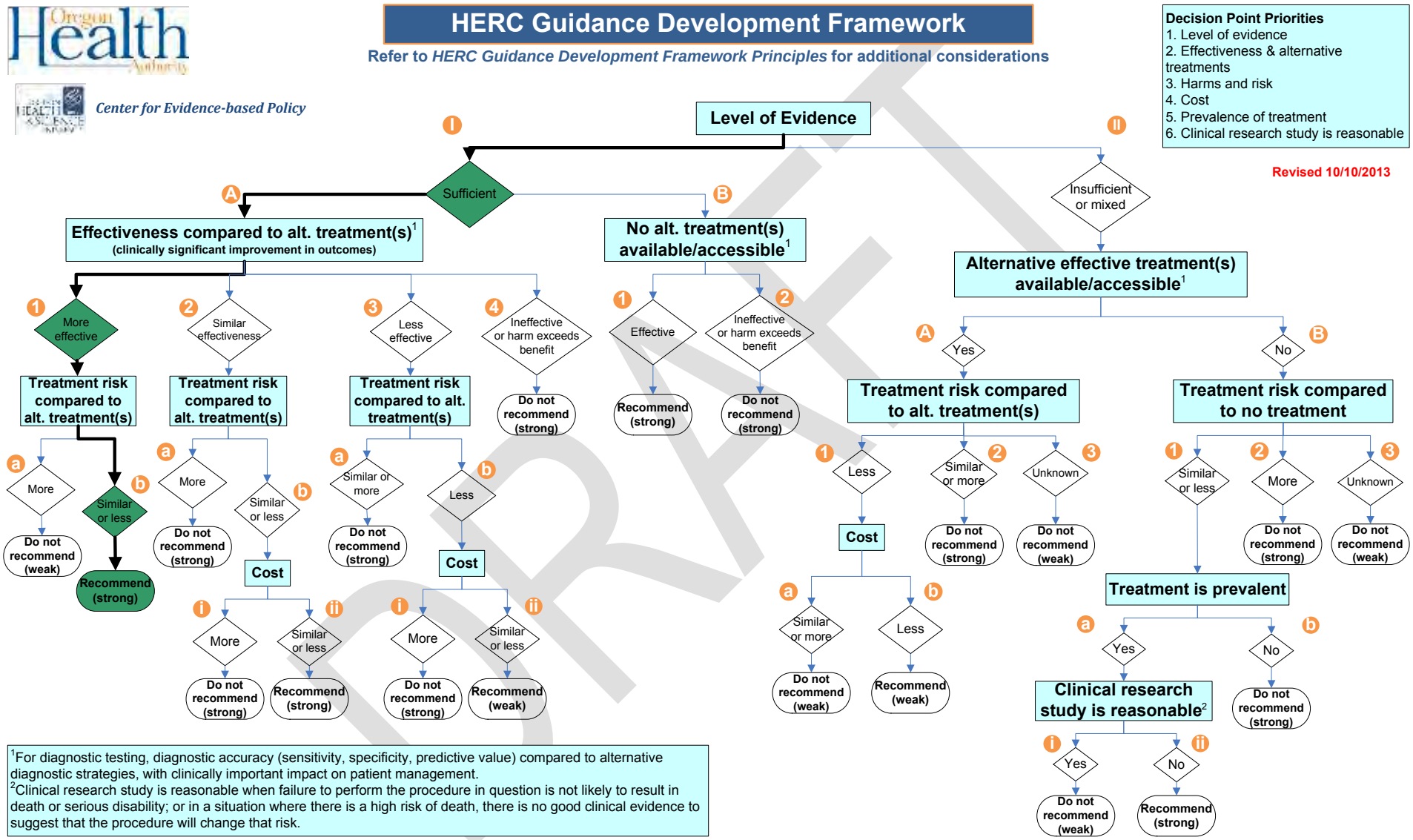
CODES	DESCRIPTION
77300	Radiation Therapy Dose Plan
77321	Special Teletx Port Plan
77332/3/4	Radiation treatment aids (simple, intermediate, complex)
77336	Continuing medical physics consultation
77370	Special medical radiation physics consultation
77417	Radiology Port Films (not seen w/ SRS/SBRT)
77421	Stereoscopic Xray Guidance (not for use with SRS/SBRT)
77427/31/99	Radiation treatment management, 5 treatments (not seen w/ SRS/SBRT)
77470	Special Radiation Treatment management (extra planning for SRS)
70010-70559	Diagnostic Radiology Head and Neck
76830/1 76856/7	US (can be used for other therapy treatment planning)
71010-71555	Diagnostic Radiology Head and Neck
72010-72295	Diagnostic Radiology Spine and Pelvic
74000-74190	Diagnostic Radiology Abdomen
74210-74363	Diagnostic Radiology Gastrointestinal Tract
74400-74485	Diagnostic Radiology Urinary Tract
74710-74775	Diagnostic Radiology Gynecological and Obstetrical
75557-75564	Diagnostic Radiology Spine and Pelvic Heart
96401-96549	Chemotherapy (can be used as a sensitizer, may indicate failure of SBRT therapy)
HCPCS Level II Codes	
None	

Note: Inclusion on this list does not guarantee coverage

DRAFT

Appendix C. HERC Guidance Development Framework

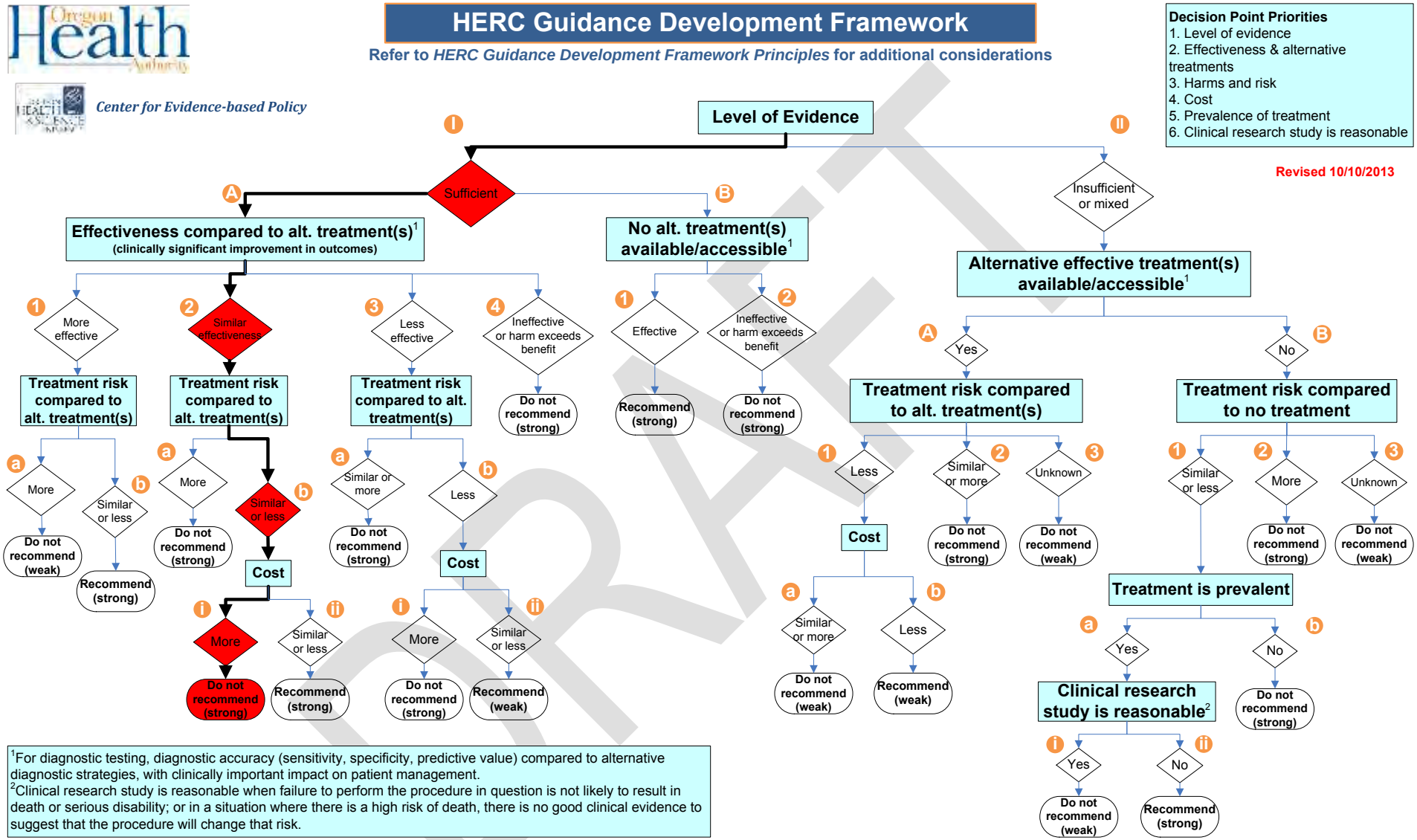
Prostate Cancer, Cervical Cancer (Consider Non-Small Cell Lung Cancer)



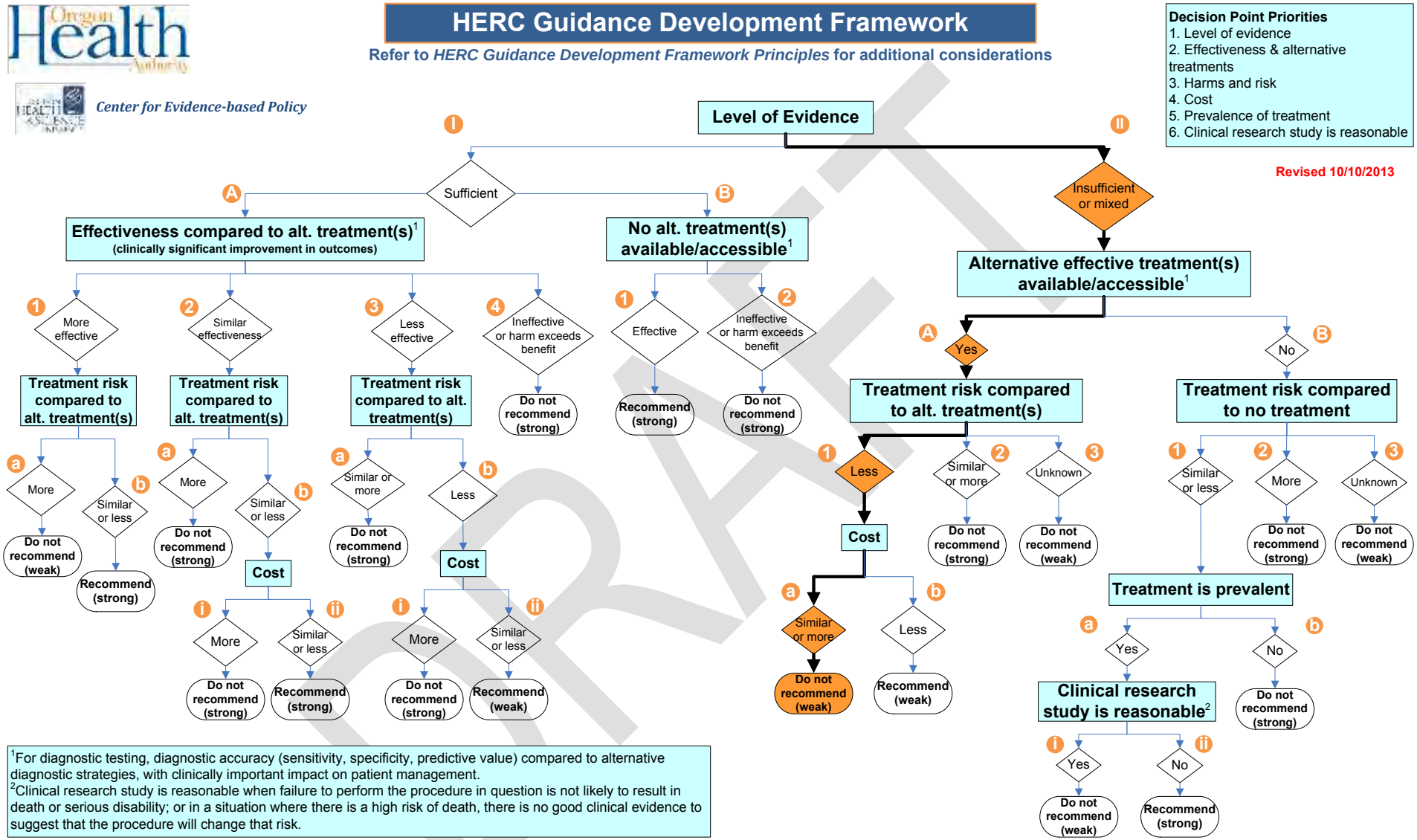
[1]For diagnostic testing, diagnostic accuracy (sensitivity, specificity, predictive value) compared to alternative diagnostic strategies, with clinically important impact on patient management.
[2]Clinical research study is reasonable when failure to perform the procedure in question is not likely to 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.

DRAFT

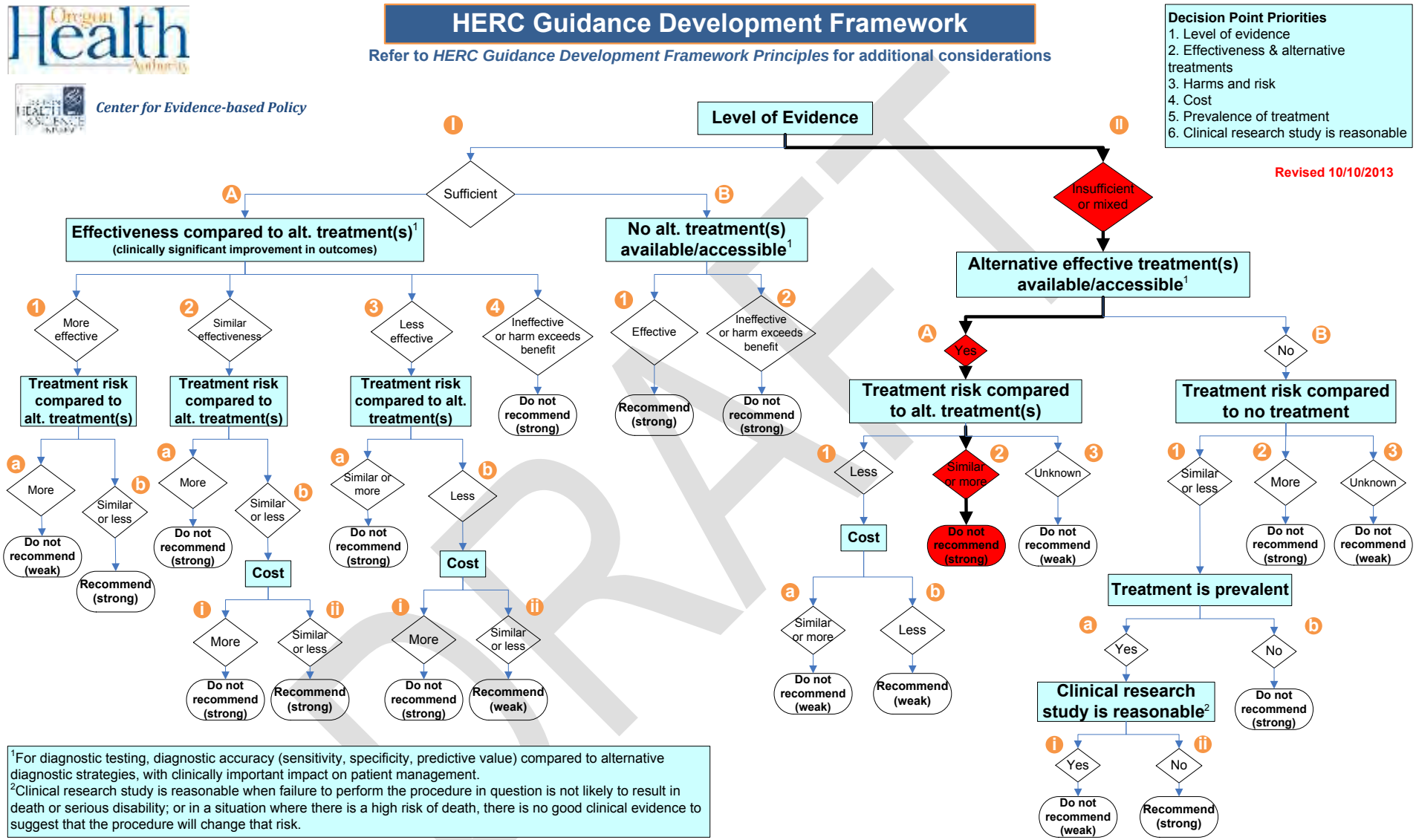
Whole Breast Radiation, Head and Neck Cancer, Thyroid Cancer (Consider Non-Small Cell Lung Cancer)



Anal Cancer, Stomach Cancer, Medulloblastoma, Para-Aortic Lymph Node Metastases

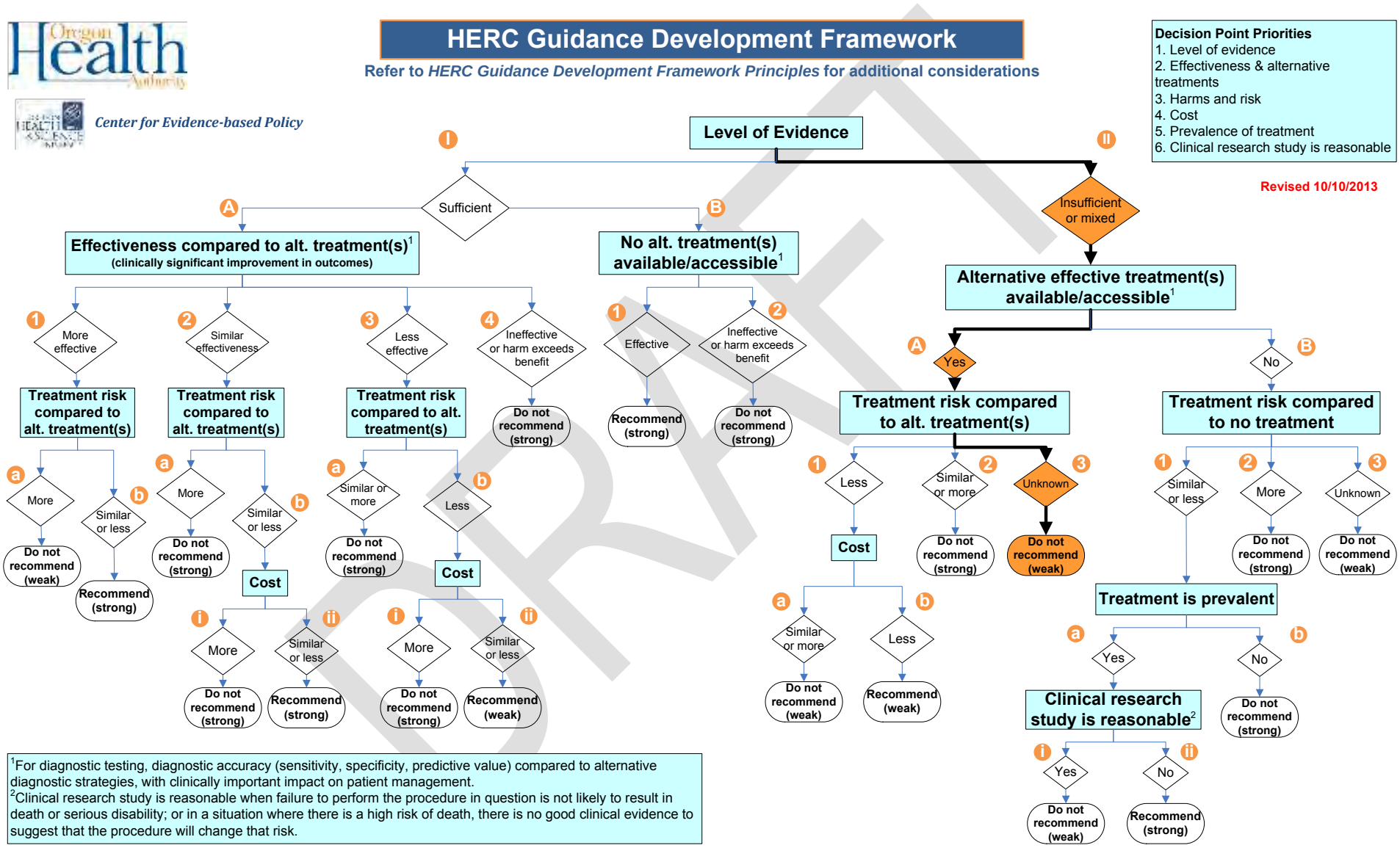


Whole Pelvis, Endometrial Cancer



DRAFT

Esophagus, Liver Pancreas, Rectum, Astrocytoma, Brain Metastases, Glioblastoma, Glioma, Meningioma, Glioma, Pituitary Adenoma, Partial Breast Radiation, Pleural Mesothelioma, Small Cell Lung Cancer, Sarcoma, Sacral Chordoma, Skin, Spinal Metastases



DRAFT

Appendix D. Guideline References from Clark 2012

Decker, R.H., Langer, C.J., Rosenzweig, K.E., Chang, J.Y., Gewanter, R.M., Ginsburg, M.E., et al. (2011). ACR Appropriateness criteria®: Postoperative adjuvant therapy in non-small cell lung cancer. *American Journal of Clinical Oncology, 34*(5), 537-544.

Gaffney, D.K., Erickson-Whittmann, B.A., Jhingran, A., Mayr, N.A., Puthawala, A.A., Cardenes, H.R., et al. (2010). *ACR appropriateness criteria®: Advanced cervical cancer*. Retrieved March 22, 2012, from http://www.acr.org/SecondaryMainMenuCategories/quality_safety/app_criteria/pdf/RO-Gyn/Advanced-Cervical-Cancer.aspx

Gewanter, R.M., Movsas, B., Rosenzweig, K.E., Chang, J.Y., Decker, R., Dubey, S., et al. Expert Panel on Radiation Oncology-Lung. (2010). *ACR Appropriateness Criteria® nonsurgical treatment for non-small-cell lung cancer: good performance status/definitive intent.* Reston, VA: American College of Radiology (ACR). Retrieved July 21, 2011, from http://www.acr.org/SecondaryMainMenuCategories/quality_safety/app_criteria/pdf/ExpertPanelonRadiationOncologyLungWorkGroup/NonsurgicalTreatmentNSCLC-GoodPerformance.aspx

Gopal, R.S., Dubey, S., Movsas, B., Rosenzweig, K.E., Chang, J.Y., Decker, R., et al. Expert Panel on Radiation Oncology-Lung. (2010). *ACR Appropriateness Criteria® induction and adjuvant therapy for N2 non-small-cell lung cancer*. Reston, VA: American College of Radiology (ACR). Retrieved July 21, 2011, from http://www.acr.org/SecondaryMainMenuCategories/quality_safety/app_criteria/pdf/ExpertPanelonRadiationOncologyLungWorkGroup/InductionandAdjuvantTherapyforN2NonSmallCellLungCancerDoc2.aspx

Lutz, S.T., Lo, S.S.M., Howell, D.D., Chang, E.L., Galanopoulos, N., Kim, E.Y., et al. (2011). *ACR Appropriateness criteria®: Non-spine bone metastases*. Retrieved March 22, 2012, from http://www.acr.org/SecondaryMainMenuCategories/quality_safety/app_criteria/pdf/ExpertPanelonRadiationOncologyBoneMETASTASESWorkGroup/BoneMETASTASESUpdateinProgressDoc1.aspx

Morgan, B.J., DeRose, P., Hsu, I.C.J., Abdel-Wahab, M., Arterbery, V.E., Ciezki, J.P., et al. (2011). ACR Appropriateness Criteria®: Definitive external beam irradiation in stage T1 and T2 prostate cancer. *American Journal of Clinical Oncology, 34*(6), 636-647.

DRAFT

National Comprehensive Cancer Network (NCCN). (2012a). *NCCN clinical practice guidelines in oncology: Anal Carcinoma. Version 2.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/anal.pdf

National Comprehensive Cancer Network (NCCN). (2012b). *NCCN clinical practice guidelines in oncology: Breast cancer. Version 1.2012.* Ft. Washington, PA: NCCN. Retrieved September 6, 2011, from http://www.nccn.org/professionals/physician_gls/pdf/breast.pdf

National Comprehensive Cancer Network (NCCN). (2012c). *NCCN clinical practice guidelines in oncology: Central nervous system cancers. Version 1.2012.* Ft. Washington, PA: NCCN. March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/cns.pdf

National Comprehensive Cancer Network (NCCN). (2012d). *NCCN clinical practice guidelines in oncology: Cervical cancers. Version 1.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/cervical.pdf

National Comprehensive Cancer Network (NCCN). (2012e). *NCCN clinical practice guidelines in oncology: Colon cancer. Version 3.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/colon.pdf

National Comprehensive Cancer Network (NCCN). (2012f). *NCCN clinical practice guidelines in oncology: Esophageal and esophagogastric junction cancers. Version 2.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/esophageal.pdf

National Comprehensive Cancer Network (NCCN). (2012g). *NCCN clinical practice guidelines in oncology: Gastric cancer. Version 2.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/gastric.pdf

National Comprehensive Cancer Network (NCCN). (2012h). *NCCN clinical practice guidelines in oncology: Head and neck cancers. Version 1.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/head-and-neck.pdf

National Comprehensive Cancer Network (NCCN). (2012i). *NCCN clinical practice guidelines in oncology: Malignant pleural mesothelioma. Version 2.2012.* Ft.

DRAFT

Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/mpm.pdf

National Comprehensive Cancer Network (NCCN). (2012j). *NCCN clinical practice guidelines in oncology: Mucosal melanoma of the head and neck. Version 1.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/mucosal.pdf

National Comprehensive Cancer Network (NCCN). (2012k). *NCCN clinical practice guidelines in oncology: Non-small cell lung cancer. Version 3.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf

National Comprehensive Cancer Network (NCCN). (2012l). *NCCN clinical practice guidelines in oncology: Prostate cancer. Version 3.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf

National Comprehensive Cancer Network (NCCN). (2012m). *NCCN clinical practice guidelines in oncology: Rectal cancer. Version 3.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/rectal.pdf

National Comprehensive Cancer Network (NCCN). (2012n). *NCCN clinical practice guidelines in oncology: Small cell lung cancer. Version 2.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/sclc.pdf

National Comprehensive Cancer Network (NCCN). (2012o). *NCCN clinical practice guidelines in oncology: Testicular cancer. Version 1.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/testicular.pdf

National Comprehensive Cancer Network (NCCN). (2012p). *NCCN clinical practice guidelines in oncology: Thymomas and thymic carcinomas. Version 2.2012.* Ft. Washington, PA: NCCN. Retrieved March 23, 2012, from http://www.nccn.org/professionals/physician_gls/pdf/thymic.pdf

Poggi, M.M., Konski, A.A., Suh, W.W., Blackstock, A.W., Herman, J.M., Hong, T.S., et al. (2010). *ACR Appropriateness criteria®: Anal cancer.* Retrieved March 22, 2012, from http://www.acr.org/SecondaryMainMenuCategories/quality_safety/app_criteria/pd

DRAFT

f/ExpertPanelonRadiationOncologyRectalanalWorkGroup/analcancerUpdateinProgressDoc1.aspx

Quon, H., Yom, S.S., Beitler, J.J., Garg, M.K., Lawson, J., McDonald, M.W., et al. (2010). *ACR Appropriateness Criteria®: Local-regional therapy for resectable oropharyngeal squamous cell carcinomas*. Retrieved March 22, 2012, from http://www.acr.org/SecondaryMainMenuCategories/quality_safety/app_criteria/pdf/ExpertPanelonRadiationOncologyHeadNeckWorkGroup/ResectableOropharyngealSquamousCellCarcinomas.aspx

Rossi, Jr., C.J., Merrick, G., Hsu, I.C., Abdel-Wahab, M., Arterbery, V.E., Ciezki, J.P., et al. Expert Panel on Radiation Oncology-Prostate. (2010). *ACR Appropriateness Criteria® postradical prostatectomy irradiation in prostate cancer.* Reston, VA: American College of Radiology (ACR). Retrieved June 20, 2011, from http://www.acr.org/SecondaryMainMenuCategories/quality_safety/app_criteria/pdf/ExpertPanelonRadiationOncologyProstateWorkGroup/PostradicalProstatectomyIrradiationinProstateCancerDoc6.aspx

Suh, W.W., Johnstone, P.A., Blackstock, A.W., Herman, J., Konski, A.A., Mohiuddin, M., et al. (2007). *ACR Appropriateness criteria®: Resectable rectal cancer*. Retrieved March 22, 2012, from http://www.acr.org/SecondaryMainMenuCategories/quality_safety/app_criteria/pdf/ExpertPanelonradiationOncologyRectalAnalWorkGroup/ResectableRectalCancerUpdateinProgressDoc4.aspx

Wolfson, A.H., Varia, M.A., Moore, D., Rao, G.G., Cardenes, H.R., Gaffney, D.K., et al. Expert Panel on Radiation Oncology-Gynecology. (2011). *American College of Radiology ACR appropriateness criteria®: Role of adjuvant therapy in the management of early stage cervical cancer*. Retrieved March 22, 2012, from http://www.acr.org/SecondaryMainMenuCategories/quality_safety/app_criteria/pdf/RO-Gyn/Early-Stage-Cervical-Cancer.aspx

DRAFT

HEALTH EVIDENCE REVIEW COMMISSION (HERC)

COVERAGE GUIDANCE: STEREOTACTIC RADIOSURGERY/ STEREOTACTIC BODY RADIATION THERAPY

DRAFT for HTAS Meeting Materials 11/25/2013 (corrected)

HERC COVERAGE GUIDANCE

For patients with brain metastases, combination therapy with stereotactic radiosurgery (SRS) and whole brain radiation therapy (WBRT)—as opposed to WBRT alone—is recommended for coverage only for those who have a single metastasis or RPA Class I status (*weak recommendation*).

For all other patients with brain metastases, SRS alone and WBRT alone (but not in combination) are recommended for coverage (*strong recommendation*).

For patients with pituitary adenoma, SRS **is recommended** / **is not recommended** for coverage (___ *recommendation*).

For patients with spinal tumors, SBRT **is recommended** / **is not recommended** for coverage (___ *recommendation*).

For patients with all other CNS and body malignancies, SRS and SBRT are not recommended for coverage (*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 Heath 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.

DRAFT

EVIDENCE SOURCE

Gerrity, M., Thielke, A., Leof, A.W., Ryan, K., Little, A., Kriz, H., & King, V. (2012). *Stereotactic radiosurgery and stereotactic body radiation therapy*. Olympia, WA: Washington State Health Care Authority Health Technology Assessment Program. Retrieved July 30, 2013, from http://www.hta.hca.wa.gov/stereotactic_radiation.html

Glossary Source
Eastern Cooperative Oncology Group. (2006). ECOG performance status. Retrieved November 12, 2013, from http://ecog.dfci.harvard.edu/general/perf_stat.html

U.S. Department of Health and Human Services, National Institutes of Health, & National Cancer Institute. (2010). *Common terminology criteria for adverse events (CTCAE). Version 4.03.* Retrieved October 25, 2013, from http://evs.nci.nih.gov/ftp1/CTCAE/About.html

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

SUMMARY OF EVIDENCE

Clinical Background

Clinical and epidemiological overview
Over the past ten years, significant advances have been made in the techniques available to deliver external beam radiation therapy (EBRT) as a treatment modality for certain cancers. The goal of these newer techniques is two-fold: to improve the targeting of radiation to the tumor to minimize damage to normal tissue and increase the dose of radiation delivered to the tumor.

Stereotactic radiosurgery (SRS) and stereotactic body radiation therapy (SBRT) have been rapidly accepted into clinical practice and are currently used for a number of cancers—most notably central nervous system, lung, colon, breast, head and neck, and prostate cancer. These are among the most prevalent cancers in the United States and have the following incidence rates:

- Brain and other nervous system cancers (6.5 per 100,000 men and women);
- Lung cancer (62.0 per 100,000 men and women);
- Colorectal cancer (47.2 per 100,000 men and women);
- Prostate cancer (156.0 per 100,000 men); and
- Oral cavity and pharynx (10.6 per 100,000 men and women).

Technology overview

Conventional EBRT, also called 2-dimensional (2DCRT) or 3-dimensional conventional radiation (3DCRT)[1], delivers photon beams of a uniform intensity and is usually given in 25 to 50 fractions (doses) delivered five days per week for 5 to 10 weeks. Stereotactic radiosurgery was initially developed in the 1950's to treat inoperable intracranial conditions. Stereotactic radiosurgery uses a single (or very limited number of) high dose(s) of radiation directed at a tumor within the central nervous system. When used outside the central nervous system, it is referred to as SBRT and is usually delivered in three to ten fractions. Multiple radiation beams are precisely targeted to the shape of the tumor from different directions instead of from a single direction or two directions. The full dose of radiation is limited to the areas of overlap of the beams and the surrounding normal tissue receives a much lower dose. Typically, SBRT is used for tumors that are considered inoperable because of their location, for patients who are high risk for surgery, or for patients who refuse surgery. Nine devices are currently approved by the Food and Drug Administration for SRS/SBRT. These devices require a minimum staff including a certified radiation oncologist, qualified medical physicist, and licensed radiation therapist to safely deliver SRS/SBRT.

Stereotactic radiosurgery and SBRT require great precision in defining the tumor and delivering the radiation because the higher doses of radiation delivered in a fraction would cause significant damage to normal tissue. Computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET) and other imaging techniques may be used to provide image guidance immediately prior to and/or during the course of radiation treatment. This approach is referred to as image guided radiation therapy. In addition, SRS and SBRT require strategies and devices that minimize patient and organ movement. These include

1) Immobilization using body cases;
2) Implantation of radiopaque markers called fiducials;
3) Real-time CT imaging systems incorporated into linear accelerators; and
4) Techniques that manage respiratory movement (e.g. abdominal compression, breath holding when the beam is on, and gating where the beam is turned on and off with the respiratory cycle).

Evidence Review

Evidence was identified that evaluated SRS and SBRT for cancers in the following anatomic locations: abdomen (anus/rectum/colon, liver, pancreas, and adrenal glands), central nervous system (astrocytoma, brain metastases, ependymoma, glioblastoma, glioma, meningioma, neurocytoma, pituitary adenoma, schwannoma), head and neck

[1] In this report 2DCRT and 3DCRT are grouped together as conventional radiation therapy (CRT) except where individual studies compare IMRT to either 2DCRT or 3DCRT. Current conventional EBRT is also referred to as CRT.

(glomus jugulare, head and neck, ocular melanoma), lung, prostate, and spine. A total of 3,034 citations were screened, of which 253 studies met criteria for inclusion in this review. Except for six randomized controlled trials (RCTs) of SRS for brain metastases and one for glioblastoma, the evidence for SRS and SBRT is based on cohort and case series studies that have substantial methodological limitations. Almost all of these studies are non-comparative, and only two focus solely on children. Thus, the risk of bias is high and estimates of the relative benefits and harms of SRS/SBRT compared to conventional EBRT are highly uncertain for most of the tumors covered in this review. Because surgery is generally not considered an option in patients undergoing SRS/SBRT, comparative evidence is limited to EBRT.

The findings from comparative studies addressing efficacy (e.g., overall survival, quality of life) and harms are summarized below by tumor. For the remainder of the tumors, the overall strength of evidence was very low and often heterogeneous. Evidence was limited either to case series or no more than one poor quality cohort study. Therefore, no general conclusions can be drawn for these tumors. In addition, even though the overall strength of evidence is low or very low, harms for a few tumors will be described because of their frequency or severity. For the remaining tumors, in addition to fatigue and general malaise, harms were mostly regional toxicities based on the location of the malignancy (e.g., radiation pneumonitis for lung, headaches or radionecrosis with brain edema for brain, erectile dysfunction for prostate) and commonly included acute and late toxicities[2] Information on cost was identified for brain metastases, lung cancer, pancreatic cancer, meningioma and tumors of the spine, and is presented in those respective sections.

Brain Metastases

For SRS + whole brain radiation therapy (WBRT) compared to WBRT alone, the overall strength of evidence is moderate for survival and tumor control. Although local tumor control is probably better, SRS+WBRT compared to WBRT alone likely has no significant difference in overall survival. Subgroup analyses from one RCT (n=333), which provides low overall strength of evidence, suggest that median survival in patients with single metastases (6.5 vs. 4.9 months, SRS+WBRT vs. WBRT, respectively, p=0.039) and patients who are recursive partitioning analysis (RPA)[3] Class 1 (11.6 vs.

[2] Toxicities are graded as follows: *Grade 1:* Mild; asymptomatic or mild symptoms; clinical or diagnostic observations only; intervention not indicated. *Grade 2:* Moderate; minimal, local or noninvasive intervention indicated; limiting age-appropriate instrumental activities of daily living (ADL) (*e.g.*, preparing meals, shopping for groceries or clothes, using the telephone, managing money, etc.). *Grade 3:* Severe or medically significant but not immediately life-threatening; hospitalization or prolongation of hospitalization indicated; disabling; limiting self-care ADL (*e.g.,* bathing, dressing and undressing, feeding self, using the toilet, taking medications, and not bedridden). *Grade 4:* Life-threatening consequences; urgent intervention indicated. *Grade 5:* Death related to adverse events.

[3] RPA is classification system related to patient prognosis. *Class 1:* Karnofsky performance status (KPS) ≥ 70, age < 65, controlled primary disease with no extracranial mets. *Class 2:* not meeting criteria for

9.6 months, SRS+WBRT vs. WBRT, respectively, p=0.045) may be better with SRS+WBRT compared to WBRT alone. Acute and late toxicities are probably not significantly different for SRS+WBRT compared to WBRT alone, based on moderate strength of evidence. Approximately 2% to 5% of patients may experience severe (Grade 3 or 4) acute and late toxicities

For SRS+WBRT compared to SRS alone, the overall strength of evidence is moderate for the outcome of overall survival and tumor control. Although local and distant tumor control is probably better, SRS+WBRT compared to SRS alone probably has no significant difference in overall survival. An overall low strength of evidence exists to suggest there is no difference in functional independence, time to worsened performance status or quality of life for SRS+WBRT compared to SRS alone. The overall strength of evidence is low for harms and indicates that severe (Grade 3 or 4) acute and late toxicities may be similar for SRS+WBRT compared to SRS alone and occur in approximately 2% to 5% of patients.

For SRS alone compared to WBRT alone, the overall strength of evidence is very low based on six cohort studies, two with historical controls, and two additional small poor quality cohort studies. These studies suggest that overall survival may be better for patients receiving SRS alone compared to WBRT alone, but the poor quality of the studies and the heterogeneity across studies limit any conclusions. For harms, toxicity rates appear to be similar for SRS alone compared to WBRT alone. For cost, the strength of evidence is very low that SRS alone is more cost-effective than WBRT alone or SRS plus WBRT based on poor quality economic evaluations.

Glioblastoma

The overall strength of the evidence is low based on one fair quality RCT that conflicts with two poor quality cohort studies. The addition of SRS to EBRT and carmustine (chemotherapy) may not affect survival in patients with newly diagnosed glioblastoma based on the results from the RCT. However, adding SRS to other treatments for glioblastoma may increase the risk of symptomatic radionecrosis requiring a second surgery, based on low overall strength of evidence.

Pituitary Adenoma

Based on one fair quality (n=125) and one poor quality (n=72) cohort study, there is a low overall quality of evidence suggesting there may be no difference in overall survival or local tumor control in patients treated with SRS instead of EBRT, but there is uncertainty regarding this conclusion. New onset hypopituitarism was lower in the SRS groups in both studies, although this was not statistically significant. For hormone secreting tumors, the median time to remission was shorter in the SRS groups in both

---

class 1 or 3. *Class 3:* KPS < 70. KPS = 70 indicates patients can take care of themselves, are out of bed more than 50% of the time, but are unable to do normal work and activities.

studies (statistically significant in one). Thirteen case series added additional information on harms. Radiation induced pituitary deficiencies ranged from 9% to 30%. Additional harms include transient cranial nerve palsy, visual deficits, temporal lobe necrosis, internal carotid artery stenosis, unilateral blindness, seizures and memory loss.

Glioma

Based on one poor quality cohort study and eight case series, the overall strength of evidence is very low for prolonged survival with salvage SRS in patients with recurrent gliomas and for harms in patients with primary and recurrent malignant gliomas. Although there is uncertainty, these studies raise concerns about radiation necrosis leading to a mass effect requiring surgery or potentially stimulating recurrence.

Meningioma

No studies addressed effectiveness. Based on 28 case series, the overall strength of evidence is very low for harms. Erythema, alopecia and post-radiation edema are common adverse effects, and those treated with gamma knife radiosurgery (GKRS) had an overall complication rate of 13% and permanent morbidity of 7%. One poor quality cost analysis compared microsurgery, linear accelerator and GKRS in the Netherlands and found microsurgery slightly more costly than the other comparators, but the poor quality of the study prevents conclusions.

Schwannoma

The overall strength of evidence for harms from SRS for schwannomas is very low, based on two poor quality cohort studies. However, about 1% of patients may develop hydrocephalus requiring a shunt (although one study suggests this is as high as 12%), 1% to 2% may develop a new malignancy, and up to 36% may develop new facial nerve dysfunction. There were no studies that compared SRS to EBRT, so relative harms are uncertain.

Ocular melanoma

The overall strength of evidence for harms from SRS for choroidal and uveal melanoma is very low. However, enucleation due to treatment side effects such as painful neovascular glaucoma may occur in 4% to 13% of patients.

Early Stage Non-Small Cell Lung Cancer

The overall strength of evidence is very low for efficacy outcomes. SBRT for non-operable Stage I non-small cell lung cancer (NSCLC) may result in 3-year overall survival rates of 50% to 60% and local control rates of 80% to 100%. The overall strength of evidence regarding harms is very low. There is uncertainty about the rate of acute and late toxicities, especially as they compared to EBRT. However, rates of greater than or equal to Grade 3 late toxicities may range 2% to 10%. In addition, the

placement of fiducial markers, when used, to help target the radiation to the tumor may cause a pneumothorax requiring chest tube placement or hospitalization in approximately 9% to 28% of patients. There is very low strength of evidence based on three poor quality economic analyses pertaining to costs. The costs of EBRT may be $50,000 to $61,000 and the costs for SBRT may be $41,000 to $57,000, with an incremental cost effectiveness ratio of $6,000 per quality adjusted life year, although there is significant uncertainty in these estimates.

Spine Tumors

The overall strength of evidence is very low based on 40 case series. Some of these studies included patients who were EBRT treatment failures. Local tumor control rates may range from 76% to 96%, and median survival may range from 11 to 22 months. Rates of pain control may range from 80% to 90%, with improvement in quality of life. However, comparative rates for EBRT are not reported. Adverse events may include fatigue, nausea, esophagitis, dysphagia, spinal fracture, paresis and myelopathy. One poor quality cost study suggests that costs/patient for SBRT may be $8,424 and the costs of EBRT may be as low as $4,999, but the overall strength of evidence is very low.

Subgroups, Cost and Cost-effectiveness

Few, if any, studies addressed patient subgroups or costs of SRS/SBRT. Except as noted above for brain metastases, there was insufficient evidence to address outcomes and harms for any subgroup for any of the tumors in this report. The cost studies done, as described above, were low quality with significant risk of bias in their estimates of effectiveness and costs. Study limitations make drawing any conclusions about cost or cost-effectiveness difficult.

Guidelines

A total of 16 guidelines and 11 ACR Appropriateness Criteria®[4] were identified that address the use of SRS and SBRT. Appropriateness Criteria® issued by ACR are considered to be a clinical decision making aid rather than a broadly applied guideline. All guidelines and Appropriateness Criteria® are summarized in the table below.

[4] In the ACR Appropriateness Criteria® scale, a score of 1 to 3 is considered "usually not appropriate", 4 to 6 is considered "may be appropriate", and 7 to 9 is considered "usually appropriate."

Summary of Guidelines and ACR Appropriateness Criteria® by Tumor Location

Malignancy	Guideline (Year) Quality	Usually Not Appropriate / Not Recommended	May be Appropriate	Usually Appropriate / Recommended
Abdomen				
Recurrent rectal cancer	Konski [ACR] 2011b Fair	In four case variants of recurrent rectal cancer presented, SBRT therapy was considered "usually not appropriate" in all cases.		
Hepatocellular carcinoma	NCCN 2012c Poor		All tumors irrespective of location may be amenable to SBRT or external-beam conformal radiation. SBRT is often used for 1-3 tumors with a cumulative diameter under 6 cm. SBRT could be considered for larger lesions, if there is at least 800 cc of uninvolved liver and liver radiation tolerance can be respected.	
Rectal cancer	NCCN 2012h Poor	In patients with a limited number of liver or lung metastases, radiotherapy can be considered in highly selected cases or in the setting of a clinical trial. Radiotherapy should not be used in the place of surgical resection.		
Colon cancer	NCCN 2012b Poor	In patients with a limited number of liver or lung metastases, radiotherapy can be considered in highly selected cases or in the setting of a clinical trial. Radiotherapy should not be used in the place of surgical resection.		

DRAFT

Malignancy	Guideline (Year) Quality	Usually Not Appropriate / Not Recommended	May be Appropriate	Usually Appropriate / Recommended
Pancreatic adenocarcinoma	NCCN 2012g Poor	No standard dose or dose per fraction has been established for SBRT; therefore, it should preferably be utilized as part of a clinical trial.		
Brain and CNS				
Melanoma	ACN 2008 Good		To improve survival, patients with limited or no extracranial disease and with favorable prognosis brain metastases can be considered for surgical resection and if unresectable, for stereotactic radiosurgery	
Brain metastases	Patel [ACR] 2011 Fair		Radiosurgery for recurrent brain metastases is a viable option if size and number permit.	
Brain metastases	Videtic [ACR] 2009 Fair	Given the finding that SRS does not increase survival of patients with two or more brain metastases, clinicians need to practice careful selection of patients for this intervention.		
Brain metastases	Suh [ACR] 2010 Fair		Since much controversy exists regarding optimal treatment for a patient with a single brain metastasis, patient participation in clinical trials is important to evaluate best treatment. For those patients who do not participate in clinical trials, the roles of surgery and SRS in improving outcomes for patients with a single lesion are evident.	
Brain metastases from	American Thyroid		EBRT (including stereotactic radiosurgery)	

DRAFT

Malignancy	Guideline (Year) Quality	Usually Not Appropriate / Not Recommended	May be Appropriate	Usually Appropriate / Recommended
thyroid cancer	Association 2009 Poor		may be indicated for brain metastases not amenable to surgery	
Brain metastases	Ammirati 2010 Poor		Re-irradiation (either WBRT and/or SRS), surgical excision or, to a lesser extent, chemotherapy, can be recommended depending on a patient's specific condition and based on the judgment of the patient's treating physician.	
Brain metastases	Tsao [ASTRO] 2012 Fair		If patient has good prognosis and brain metastasis ≤ 3-4 cm. For multiple brain metastases, patients with good prognosis and all metastases ≤ 3-4cm.	
Brain metastases	IRSA 2008 Poor			The available data indicate that SRS and open surgical resection (where feasible) are both excellent treatment options for patients with solitary brain metastases. Stereotactic radiosurgery is an effective treatment for patients with multiple brain metastases
Low grade glioma	NCCN 2012a Poor	SRS has not been established to have a role in the management		

DRAFT

Malignancy	Guideline (Year) Quality	Usually Not Appropriate / Not Recommended	May be Appropriate	Usually Appropriate / Recommended
		of low grade gliomas. Phase I trials using SRS do not support its role as initial treatment.		
Meningioma	NCCN 2012a Poor		WHO grade 1 meningiomas may also be treated with stereotactic radiosurgery doses of 12-14 Gy in a single fraction when appropriate.	
Brain metastases	NCCN 2012a Poor			Recommended maximum marginal doses of 24, 18, or 15 Gy according to tumor volume is recommended.
Metastatic Spine	NCCN 2012a Poor		Doses to vertebral body metastases will depend on patient's performance status and primary histology. In selected cases, or recurrences after previous radiation, stereotactic radiotherapy is appropriate.	
Brain metastases from thyroid cancer	NCCN 2012j Poor			For solitary CNS lesions, either neurosurgical resection or stereotactic radiosurgery is preferred.
Head and Neck				
Recurrent head and neck	McDonald[ACR] 2010 Fair		SBRT therapy "may be appropriate" in one of five cases. SBRT was not considered in the treatment for the remaining four cases.	
Lung				
Stage I/II NSCLC	Scott [ACCP]		Other local therapies such as **stereotactic radiation** or radiofrequency ablation	

DRAFT

Malignancy	Guideline (Year) Quality	Usually Not Appropriate / Not Recommended	May be Appropriate	Usually Appropriate / Recommended
	2007 Fair		may be appropriate for **patients who are medically inoperable**. The use of these techniques in patients who are surgical candidates should not occur outside of the context of a clinical research study.	
Stage I NSCLC	Gewanter [ACR] 2010 Fair		Emerging institutional data suggest that central early-stage lung lesions can be treated safely with lower doses per fraction	
Stage I NSLCL	Rosenzweig [ACR] 2008 Fair	Currently **extracranial stereotactic body radiotherapy (SBRT)** is being examined as an alternative to conventionally fractionated radiotherapy in patients with inoperable stage I disease		
Stage I	NCCN 2012f Poor		Recommended for patients who are medically inoperable and is also an appropriate option for many older patients	
Prostate				
	Morgan [ACR] 2011 Fair	The use of hypofractionation in general and a stereotactic approach looks very promising, but more robust studies with longer follow-up clearly are needed.		
Other Cancers/ Multiple Sites				
Bone metastases	Janjan [ACR] 2008 Fair	SBRT therapy was considered to be "usually not appropriate" in seven of 8 cases. SBRT was not		

DRAFT

Malignancy	Guideline (Year) Quality	Usually Not Appropriate / Not Recommended	May be Appropriate	Usually Appropriate / Recommended
		considered in the treatment for the remaining case.		
Non-spine bone metastases	Lutz [ACR] 2011 Fair	SBRT therapy was considered to be "usually not appropriate" in four of five cases. SBRT was not considered in the treatment for the remaining case.		
Soft tissue sarcoma	NCCN 2012i Poor		Patients can also receive stereotactic radiosurgery or chemotherapy as an alternate method for control of metastatic lesions. Many different issues are factored into this decision (e.g., patient performance status, patient preferences, specific clinical problems from the metastases, treatment availability), and specific details are best left to clinical judgment.	

[Evidence Source]

Evidence Summary

In patients with brain metastases, there is a moderate level of evidence for SRS+WBRT compared to WBRT alone for survival and tumor control. Local tumor control is probably better for SRS+WBRT compared to WBRT alone, but likely has no significant difference in overall survival. One RCT (low strength of evidence) suggests longer median survival in patients with single metastases and patients who are RPA Class 1 (a prognostic measure based on age, performance status and tumor control). There is a moderate level of evidence for SRS+WBRT compared to SRS alone for the outcome of overall survival and tumor control. Although local and distant tumor control is probably better, SRS+WBRT compared to SRS alone probably has no significant difference in overall survival. There is a low level of evidence that there is no difference in overall survival or local tumor control in patients with pituitary adenoma treated with SRS instead of EBRT.

DRAFT

Similarly, there is a low level of evidence that in patients with newly diagnosed glioblastoma, the addition of SRS to EBRT and chemotherapy does not affect survival. There is an insufficient level of evidence for all other tumors and comparisons.

DRAFT

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	Balance between desirable and undesirable effects	Quality of evidence*	Resource Allocation	Values and preferences	Coverage Recommendation
SRS + WBRT vs. WBRT alone for brain mets	Better local control, no diff in overall survival or harms	Moderate	Addition of SRS would be more costly than WBRT alone	Moderate to high variability	For patients with brain metastases, combination therapy with SRS and WBRT (as opposed to WBRT alone) is recommended for coverage only for those who have a single metastasis or RPA Class I status (*weak recommendation*) For all other patients with brain metastases, WBRT or SRS alone is recommended for coverage (*strong recommendation*)
SRS + WBRT vs. SRS alone for brain mets	Better local and distant control, no diff in quality of life, overall survival or harms	Low/ moderate	SRS alone may be more cost effective	Moderate to high variability	
SRS in addition to EBRT and chemo for glioblastoma	No diff in survival, increased harm	Low	Addition of SRS would be more costly	Moderate to high variability	For patients with glioblastoma, SRS in addition to chemo and EBRT is not recommended for coverage (strong recommendation)
SRS vs. EBRT for pituitary adenoma	No diff in overall survival, local control	Low	Unknown	Moderate to high variability	To be determined
SBRT for spinal tumors	Unknown	Very low	SBRT may be more costly	Moderate to high variability	To be determined
All other tumors	Unknown	Insufficient	Unknown	Moderate to high variability	For patients with all other CNS and body malignancies, SRS and SBRT

DRAFT

Indication	Balance between desirable and undesirable effects	Quality of evidence*	Resource Allocation	Values and preferences	Coverage Recommendation
					are not recommended for coverage (*weak recommendation*).

*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

DRAFT

POLICY LANDSCAPE

No quality measures were identified when searching the National Quality Measures Clearinghouse.

Medicare has not issued a national coverage determination for SRS/SBRT. Coverage decisions are therefore issued by regional Medicare contractors through Local Coverage Determinations (LCDs). Two Medicare LCDs that cover Washington were reported in the source report: one addressing SBRT (L28366 [2011]), and another addressing SRS and stereotactic radiotherapy[5] (SRT) (L30318 [2011]) (CMS 2011b, 2011c). The Medicare LCDs identify coverage of SBRT for the following indications.

SBRT: LCD 28366 (2011) states that SBRT is covered for primary and metastatic tumors of the lung, liver, kidney or pancreas when the following criteria are met:

- Patient's medical condition justified aggressive treatment;
- Other forms of radiotherapy or focal therapy (including but not limited to EBRT and IMRT) cannot be as safely or effectively utilized;
- The tumor can be completely targeted with acceptable risk to surrounding critical structures;
- For germ cell or lymphoma, effective chemotherapy regimens have been exhausted or are not otherwise feasible; and
- When other forms of focal therapy cannot be as safely or effectively used.

Coverage is possible for other lesions with documented necessity. Coverage for SBRT is not covered for the following conditions and circumstances:

- Treatment is unlikely to result in clinical cancer control and/or functional improvement;
- When there is wide-spread cerebral or extra-cranial metastases; or
- Patient has poor performance status.

For prostate cancer, SBRT is covered as monotherpay for low and intermediate risk prostate cancer when:

- Patient's medical condition justified aggressive treatment;
- Other forms of radiotherapy or focal therapy (including but not limited to EBRT and IMRT) cannot be as safely or effectively utilized; and
- The tumor can be completely targeted with acceptable risk to surrounding critical structures;

Lesions of other sites (bone, breast, uterus, ovary, and other internal organs) are generally not covered, but may be in cases of recurrence after conventional radiation modalities.

[5] In stereotactic radiotherapy, radiation is delivered in multiple fractions (2-5) at a somewhat lower dose than SRS, to a larger area

SRS: LCD 30318 (2011) states that intracranial lesions are covered under the following conditions:

- The lesion(s) has an image-distinct margin; and
- Karnofsky performance scale > 50% or ECOG[6] ≤ 2.

Specific indications include neuromas of the cranial nerves, and unresectable/residual meningioma where surgery is not appropriate. Metastatic brain lesions are covered when patients should have otherwise stable disease, margins are distinct, and treatment is for less than five lesions. SRS is also covered as a boost treatment for larger lesions treated with EBRT or surgery, acoustic neuromas, pituitary adenomas, craniopharyngiomas, and glomus jugulare tumors.

SRT: LCD 30318 (2011) states that SRT is considered medically necessary for the treatment of tumors in hard to reach locations, unusual shapes, and close proximity to vital structure. Specific indications include:

- Benign lesions (e.g., pituitary adenoma, vestibular schwannoma, meningioma);
- Benign neoplasms previously treated with conventional radiotherapy; and
- Malignant lesions (lesions less than 5mm of the optic nerves or chiasms, recurrent malignant gliomas, brain metastasis, base of skull, recurring malignancies in head and neck cancers, such as cancer of the tonsil, larynx, tongue, sinus, and mouth).

Aetna

Coverage for SBRT is limited to localized malignant conditions where highly precise application is required. This includes lung or liver metastases not amenable to surgery, medically inoperable early stage lung cancer, primary liver cancer not amenable to surgery, spinal and para-spinous tumors, though this is not an exhaustive list. The use of SRS is considered medically necessary for the treatment of benign tumors considered unresectable due to deep intracranial location or if the patient cannot tolerate surgery. Brain malignancies are also covered, both primary and metastatic. When the coverage criteria for SRS is met, SRT is considered medically necessary for tumors with such proximity to vital structures that even very accurate high-dose single fraction SRS could not be tolerated (Aetna 2011).

COMMITTEE DELIBERATIONS – HTAS

COMMITTEE DELIBERATIONS – VBBS

[6] ECOG is the Eastern Cooperative Oncology Group. They have developed a performance status rating scale, in which a score of < 2 indicates that a patient is capable only of limited self-care and is confined to a bed or chair more than 50% of waking hours.

DRAFT

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 of evidence across studies for the treatment/outcome

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.

DRAFT

Appendix B. Applicable Codes

CODES	DESCRIPTION
ICD-9 Diagnosis Codes	
140	Malignant neoplasm of lip
141	Malignant neoplasm of tongue
142	Malignant neoplasm of major salivary glands
143	Malignant neoplasm of gum
144	Malignant neoplasm of floor of mouth
145	Malignant neoplasm of other and unspecified parts of mouth
146	Malignant neoplasm of oropharynx
147	Malignant neoplasm of nasopharynx
148	Malignant neoplasm of hypopharynx
149	Malignant neoplasm of other and ill-defined sites within the lip oral cavity and pharynx
155.0	Malignant neoplasm of liver, primary
155.2	Malignant neoplasm of liver, not specified as primary or secondary
157	Malignant neoplasm of pancreas
160	Malignant neoplasm of nasal cavities middle ear and accessory sinuses
161	Malignant neoplasm of larynx
162	Malignant neoplasm of trachea bronchus and lung
163	Malignant neoplasm of pleura
185	Malignant neoplasm of prostate
190	Malignant neoplasm of eye
191	Malignant neoplasm of brain
192	Malignant neoplasm of other and unspecified parts of nervous system
197.5	Secondary malignant neoplasm of large intestine and rectum
197.7	Malignant neoplasm of liver, secondary
197.8	Secondary malignant neoplasm of other digestive organs and spleen
198.3	Secondary malignant neoplasm of brain and spinal cord
198.7	Secondary malignant neoplasm of adrenal gland
198.89	Secondary malignant neoplasm of other specified sites
225	Benign neoplasm of brain and other parts of nervous system
227.3	Benign neoplasm of pituitary gland and craniopharyngeal duct
ICD-10 Diagnosis Codes	
C00-14	Malignant neoplasms of lip, oral cavity and pharynx
C18	Malignant neoplasm of colon
C19	Malignant neoplasm of rectosigmoid junction
C20	Malignant neoplasm of rectum
C21	Malignant neoplasm of anus and anal canal
C22	Malignant neoplasm of liver and intrahepatic bile ducts
C25	Malignant neoplasm of pancreas
C30	Malignant neoplasm of nasal cavity and middle ear
C31	Malignant neoplasm of accessory sinuses
C32	Malignant neoplasm of larynx
C33	Malignant neoplasm of trachea
C34	Malignant neoplasm of bronchus and lung
C61	Malignant neoplasm of prostate
C69-72	Malignant neoplasms of eye, brain and other parts of central nervous system

DRAFT

CODES	DESCRIPTION
C74	Malignant neoplasm of adrenal gland
C76	Malignant neoplasm of other and ill-defined sites
C79	Secondary malignant neoplasm of other and unspecified sites
D33	Benign neoplasm of brain and other parts of central nervous system
D35	Benign neoplasm of other and unspecified endocrine glands
ICD-9 Volume 3 (Procedure Codes)	
92.3	Stereotactic radiosurgery
93.59	Other immobilization, pressure, and attention to wound
CPT Codes	
61795	Stereotactic computer assisted volumetric (navigational) procedure, intracranial, extracranial, or spinal
61796	Stereotactic radiosurgery (particle beam, gamma ray, or linear accelerator); 1 simple cranial lesion
61797	Each additional cranial lesions, simple
61798	Complex cranial lesion
61799	Each additional cranial lesion, complex
61800	Application of stereotactic headframe for stereotactic radiosurgery
63620 /1	Stereotactic radiosurgery (particle beam, gamma ray, or linear accelerator); 1 spinal lesion (63621 for each add'l spinal lesion)
76830 /1 76856 /7	US (can be used for other therapy treatment planning
77011	Computed tomography guidance for stereotactic localization
77371	Radiation treatment delivery, stereotactic radiosurgery (SRS), complete course of treatment of cranial lesions(s) consisting of 1 session; multi-source Cobalt 60 based
77372	As 77371, but linear accelerator based
77373	Stereotactic body radiation therapy, treatment delivery, per fraction to 1 or more lesions, including image guidance, entire course not to exceed 5 fractions
77427 /31 /99	Radiation Treatment Management
77432	Stereotactic radiation treatment management of cranial lesions(s) (complete course of treatment consisting of 1 session)
77435	Stereotactic body radiation therapy, tx management, per tx course, to 1 or more lesions, w/ image guidance, max 5 fractions
G0339	Image-guided robotic linear accelerator-based stereotactic radiosurgery, complete course of therapy in one session or first session of fractionated treatment
G0340	Image-guided robotic linear accelerator-based stereotactic radiosurgery, delivery including collimator changes and custom plugging, fractionated treatment, all lesions, per session, second through fifth sessions, maximum 5 sessions/ course of treatment
G0173	Linear accelerator based stereotactic radiosurgery, complete course of therapy in one session
20665	Removal of fixation device

DRAFT

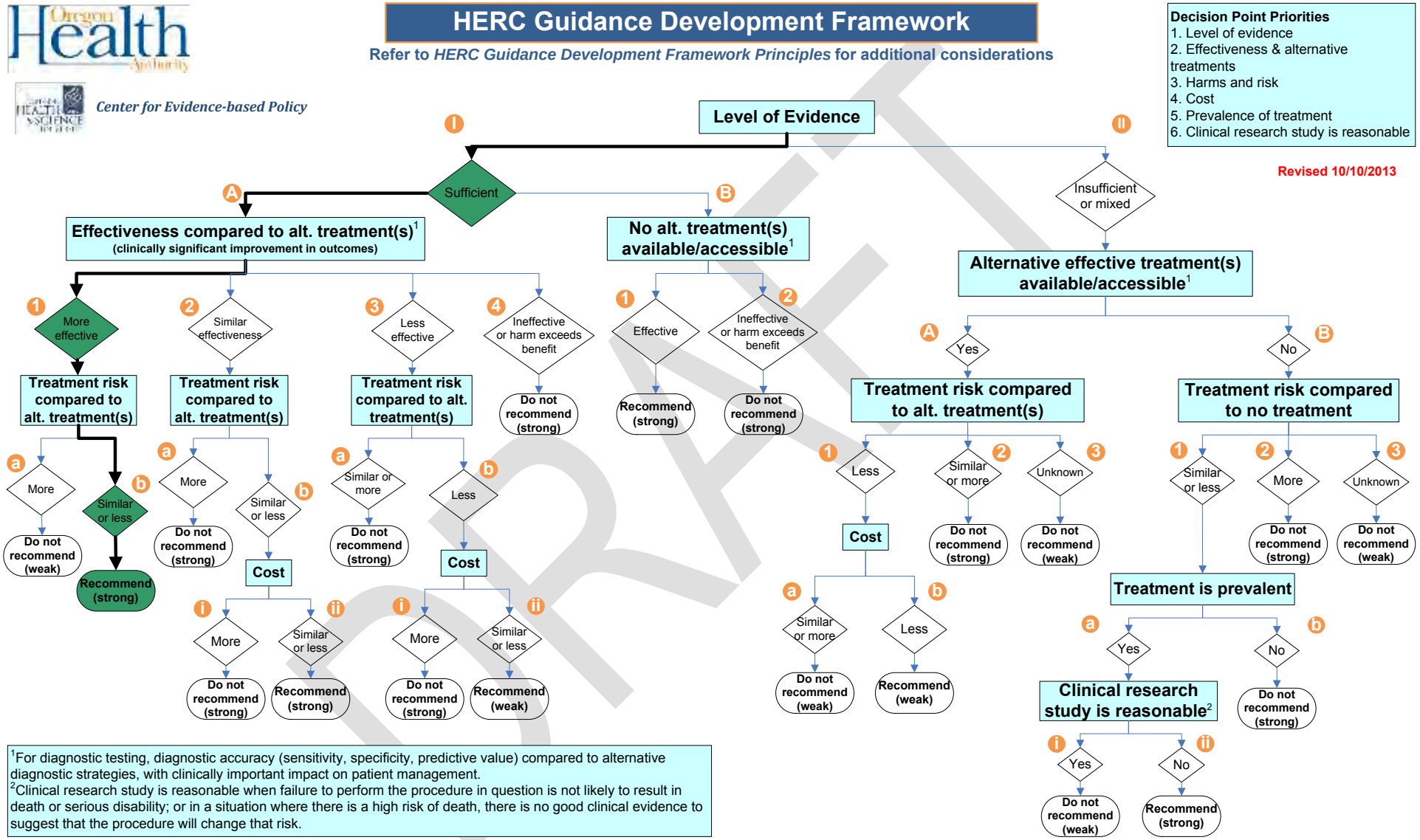
CODES	DESCRIPTION
77014	CT guidance for placement of radiation therapy fields
77261 /2 /3	Radiation Therapy Planning, simple, intermediate, complex
77280 /85 77290 /95 /99	Set radiation therapy field, simple, intermediate, complex (0) or 3 dimensional (5)
77300	Radiation Therapy Dose Plan
77321	Special Teletx Port Plan
77332 /3 /4	Radiation treatment aids (simple, intermediate, complex)
77336	Continuing medical physics consultation
77370	Special medical radiation physics consultation
77470	Special Radiation Treatment management (extra planning for SRS)
70551 /2 /3	MRI Brain
70010-70559	Diagnostic Radiology Head and Neck
71010-71555	Diagnostic Radiology Head and Neck
72010-72295	Diagnostic Radiology Spine and Pelvic
74000-74190	Diagnostic Radiology Abdomen
74210-74363	Diagnostic Radiology Gastrointestinal Tract
74400-74485	Diagnostic Radiology Urinary Tract
74710-74775	Diagnostic Radiology Gynecological and Obstetrical
75557-75564	Diagnostic Radiology Spine and Pelvic Heart
HCPCS Level II Codes	
None	

Note: Inclusion on this list does not guarantee coverage

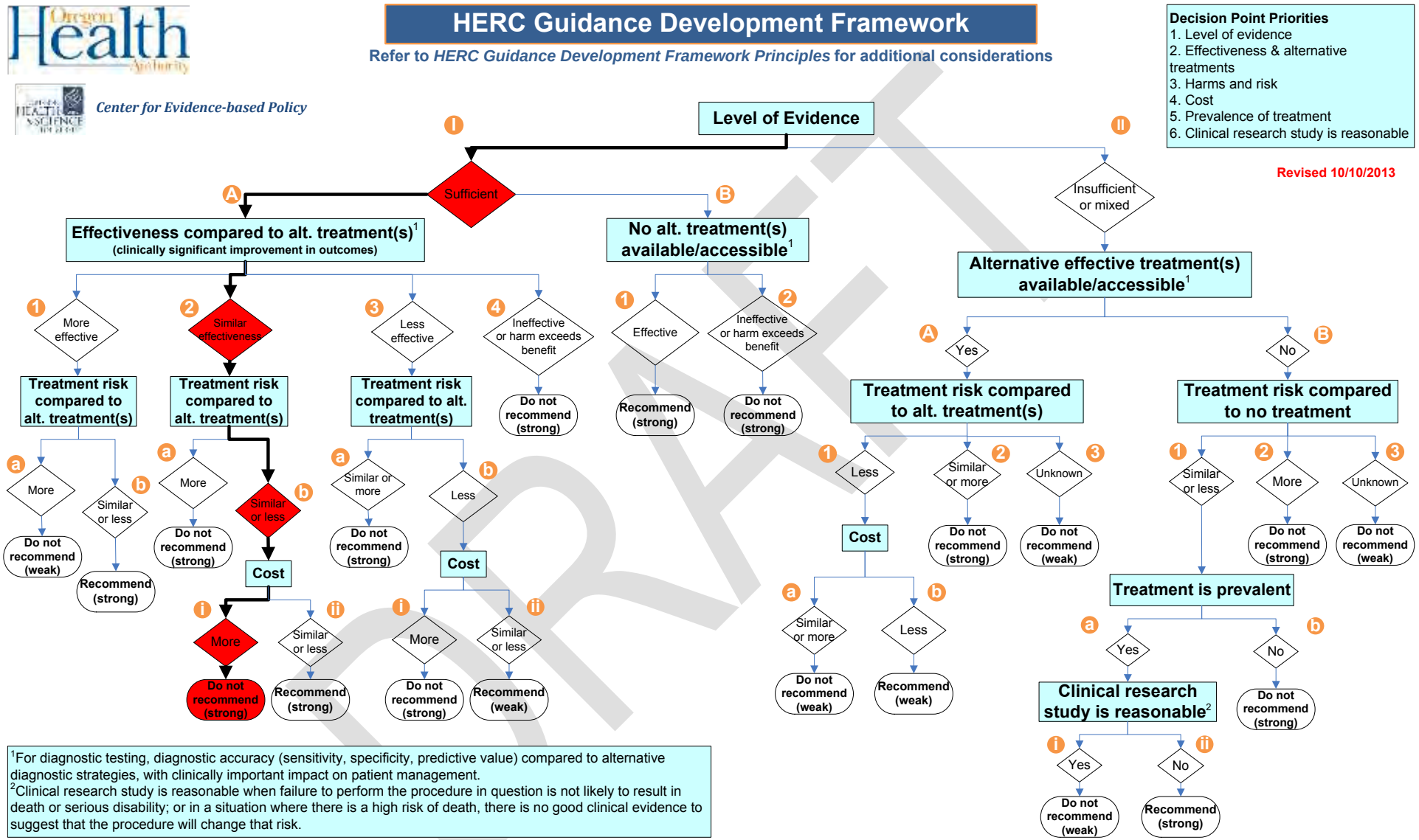
DRAFT

Appendix C. HERC Guidance Development Framework

Brain Metastases, SRS plus WBRT, Single Metastasis or RPA Class 1



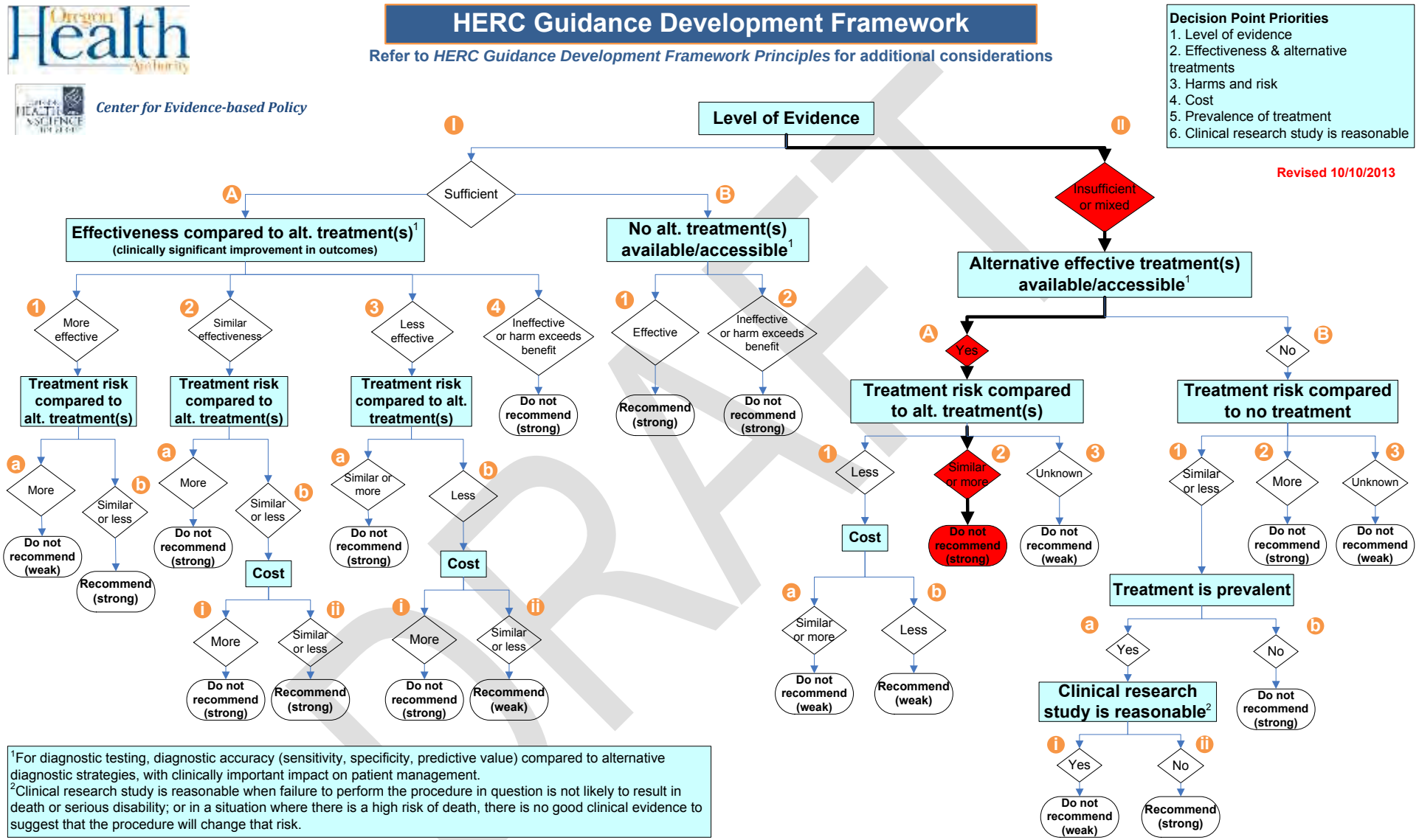
Brain Metastases, SRS Alone



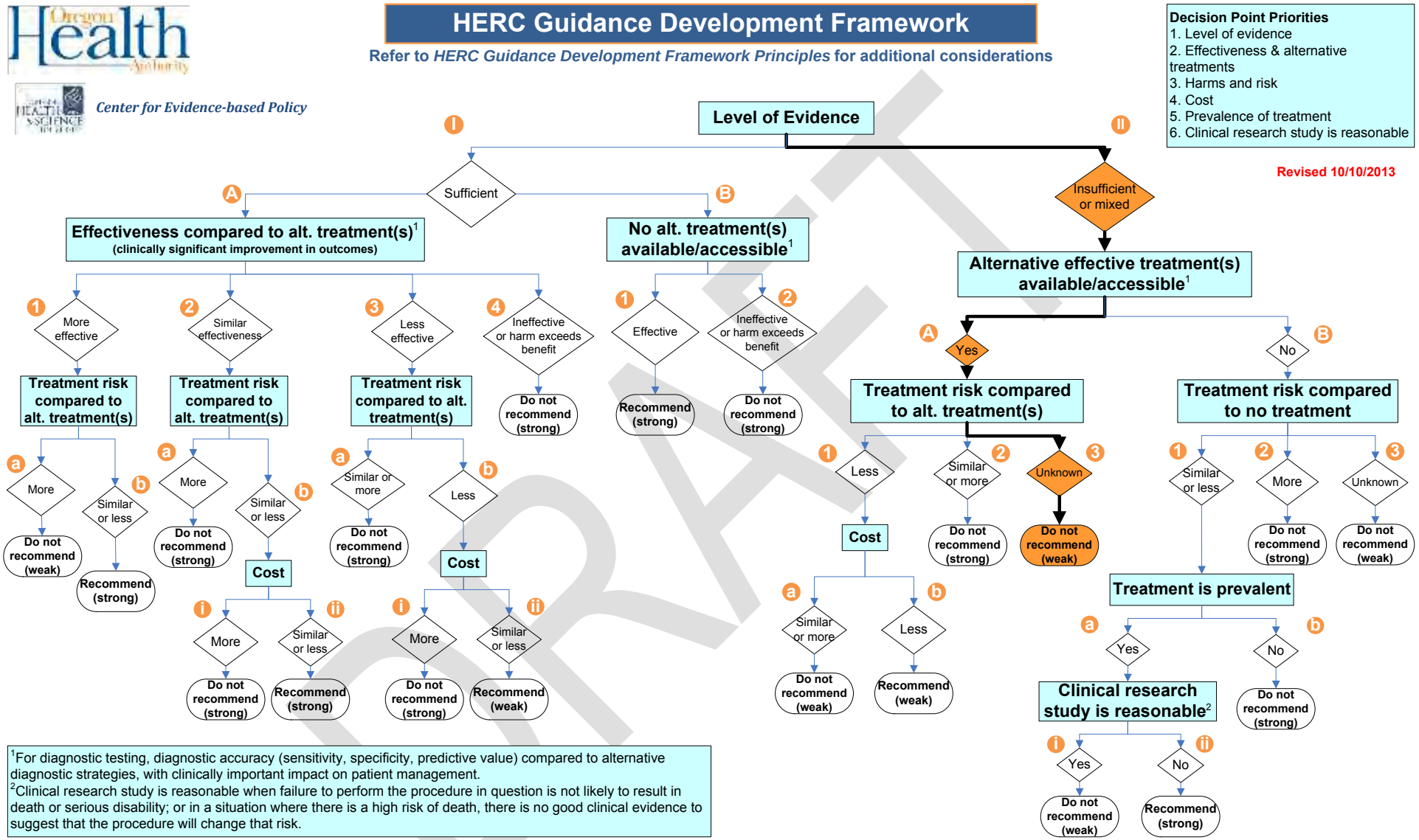
[1]For diagnostic testing, diagnostic accuracy (sensitivity, specificity, predictive value) compared to alternative diagnostic strategies, with clinically important impact on patient management.
[2]Clinical research study is reasonable when failure to perform the procedure in question is not likely to 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.

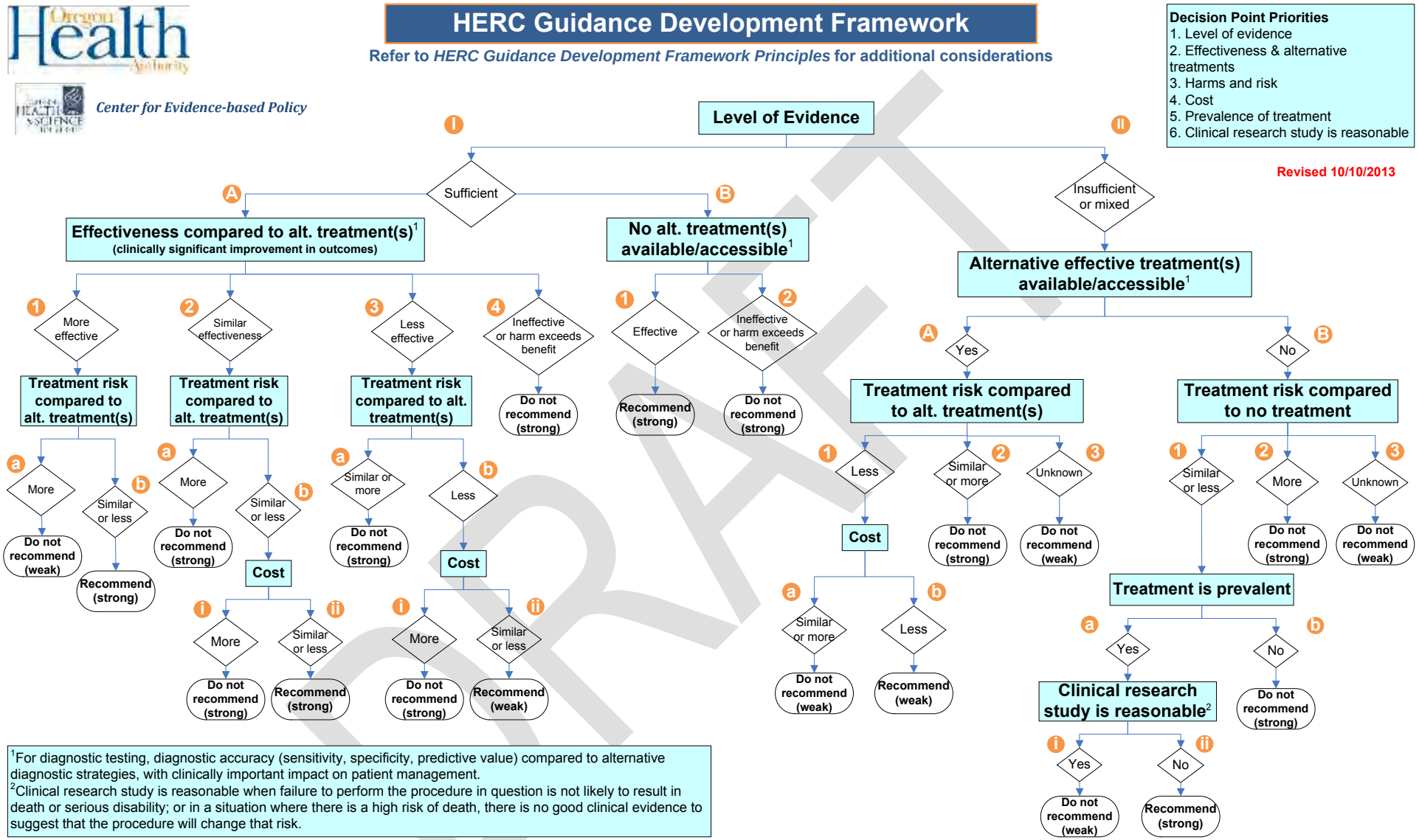
DRAFT

Glioblastoma



All Other Tumors



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

Pituitary Adenoma; Spinal Tumors – To Be Determined