



**Health Evidence Review
Commission's
Health Technology Assessment
Subcommittee**

September 23, 2013

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

AGENDA

HEALTH TECHNOLOGY ASSESSMENT SUBCOMMITTEE (HTAS)

Meridian Park Hospital Health
Education Center, Room 117B&C
19300 SW 65th Avenue, Tualatin, OR 97062
September 23, 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 June minutes	Alissa Craft	X
3	1:10 PM	Review Public Comment on draft coverage guidance on Hip Surgery Procedures For Femoroacetabular Impingement Syndrome	Alison Little Wally Shaffer	X
4	1:50 PM	Review Public Comment on draft coverage guidance on Upper Endoscopy For Gastroesophageal Reflux Disease (GERD) And Dyspepsia Symptoms	Alison Little Wally Shaffer	X
5	2:20 PM	Review Public Comment on draft coverage guidance on Osteoporosis Screening And Monitoring By Dual-Energy X-Ray Absorptiometry (DXA)	Alison Little Wally Shaffer	X
6	3:10 PM	Review draft coverage guidance on Treatment of Sleep Apnea in Adults	Wally Shaffer	X
6	3:45 PM	General Public Comment		
7	3:50 PM	Next Topics	Alissa Craft	
8	3:55 PM	Confirm next meeting: November 25, 2013 at Meridian Park Health Ed. Center, 1-4 p.m.	Alissa Craft	
8	4:00 PM	Adjournment	Alissa Craft	

MINUTES

Health Technology Assessment Subcommittee
Meridian Park Community Health Education Center
19300 SW 65th Avenue, Tualatin, OR
June 24, 2013
1:00-4:00pm

Members Present: Alissa Craft, DO, MBA, Chair; James MacKay, MD, Vice-Chair (via phone); Gerald Ahmann, MD; George Waldmann, MD; Tracy Muday MD; Timothy Keenen, MD (arrived 3:37 PM).

Members Absent: None.

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

Ad Hoc Experts: FAI Surgery - Andrea Herzka, MD (OHSU); GERD - Drew Schembre, MD (Swedish Medical Center, Seattle, WA); DXA - E. Michael Lewiecki, MD (New Mexico Clinical Research & Osteoporosis Center).

Also Attending: Alison Little, MD (CEBP); Shannon Vandergriff (CEBP); Joanie Cosgrove (LHNW); Bill Struyk (Johnson and Johnson); Denise Taray (DMAP).

1. CALL TO ORDER

Alissa Craft called the meeting of the Health Technology Assessment Subcommittee (HTAS) to order at 1:00 pm. She introduced new member Tracy Muday, MD.

2. MINUTES REVIEW

No changes were made to the April 22, 2013 minutes.

Action: Minutes approved as submitted, **5-0 (Absent: Keenen)**.

3. STAFF REPORT

Coffman reviewed the updated Guidance Development Framework, adding pathway II.3.B for situations where there is insufficient or mixed evidence for the treatment and its relative risk is unknown.

Coffman reported on the status of the coverage guidance on the treatment of sleep apnea in adults. HERC decided that if the algorithm resulted in a weak recommendation to not cover the surgery, as they suspected, it would not need to go through the entire process. As that is the case, it is being returned to VbBS for reconsideration for application to the Prioritized List.

Shaffer noted HTAS may wish to keep track of this. At the moment HERC wants to make a change related to surgery. However the OHP medical directors have raised concerns about the AHI levels in the guidance not being directly supported by the evidence. It may come back to this subcommittee, particularly since the guidance was developed before HERC adopted the algorithm and modified GRADE methodology.

Coffman noted that all other coverage guidances forwarded to HERC by HTAS were accepted without changes. VbBS considered HTAS developed coverage guidances on diagnosis of sleep apnea, continuous blood glucose monitoring, self monitoring of blood glucose, and carotid endarterectomy. No changes to the CG are being recommended for the latter two which have yet to be considered by HERC, but the guidelines are somewhat different than the guidances. For example, 50 blood glucose test strips per 90 days are recommended to be covered instead of 100 for those type 2 diabetics not requiring multiple daily insulin injections who meet certain criteria.

4. REVIEW OF NEW DRAFT COVERAGE GUIDANCES

A. Femeroacitabular Impingement (FAI) Syndrome

Wally Shaffer reviewed the draft coverage guidance included in the meeting materials. There are no significant harms of the treatment beyond those typical of surgery. The limitations of the evidence were summarized pretty well in the summary accompanying the NICE guideline review. The Washington HTA found insufficient to recommend coverage but NICE came to a different conclusion, establishing coverage under a registry approach.

Andrea Herzka, the appointed expert on this topic, said that very few surgeons are performing surgery for FAI syndrome prophylactically. Surgical patients typically have debilitating pain, have failed a trial of conservative management, and have functional limitations in addition to having a labral tear and FAI morphology shown radiologically. There is some controversy in teenagers as to whether steroid injection into the joint and extensive PT is better than proceeding straight to surgery.

She considers this treatment similar to ACL surgery in that they are being fixed all the time. The patient is not told they won't get arthritis (although it may be prevented); rather the goal is to alleviate pain and restore function.

Herzka shared slides she recently presented at Grand Rounds. The patient is typically between the ages of 16-50, with symptoms that start out as mild then progress (difficulty getting in and out of the car very common). There isn't great literature on the natural history of the disease. Initially there was an assumption that all patients with labral tear are destined for doom, but that doesn't appear to be the case.

Herzka indicated that radiologists nationwide can detect the FAI morphology pretty well on x-ray, MRI, and MRI arthrogram. About 20% of men and 2% of women of European decent have the cam type and 17% of women have the pincer type. Most common is a combination of cam/pincer types, especially in patients over 35. In Asian populations prevalence is much lower. Herzka said for some we shouldn't receive a trial of 3 months of PT because when

the cartilage gets to a certain stage, the operation becomes more difficult, but there is no consensus.

Herzka reviewed the history of the different surgical approaches through a series of studies. Initially Ganz did an open hip dislocation with complete removal of the labrum. Later he sewed it back down and the patients did much better, which is supported by studies by Philipon and Larson, who performed the procedure arthroscopically. Clohisy used a mini-open arthroplasty. This is level 4 cohort retrospective study, but all of the studies consistently show good results in disabled populations where everything else has failed. She said that because there are not other good options, commercial insurers have created a pathway to coverage through medical necessity criteria. The data is not ideal and, to a point, insufficient; there is no data for PT alone. All we can say is that the vast majority of surgical patients have failed PT.

Initially the Modified Harris Hip Score was used to measure outcomes, but more recently used the Hip Outcome Score (HOS), while others have used a Return-to-Sports Score. Return to same level of sport has been tracked and 70-90%. There is no comparison versus what percent just do, but if you're getting by and you are satisfied with conservative management, surgery shouldn't be done.

Jim Mackay asked what the natural history is. Mackay indicated he practiced for many years and doesn't recall seeing people with these complaints. Herzka said there was a study of contralateral hips for those only having surgery on one hip. Those patients who got arthritis on surgical hip also got it in the other hip. In the Villar study, 43 hips with mild to moderate arthritis and a with pistol grip deformity, 2/3 went on to develop arthritis and 1/3 didn't. So not all those with labral tears are predisposed to arthritis and it is not known why some are and some aren't. So there is no great natural history. There is no study looking at the population who had symptoms for 7+ years and simply modified their lifestyle and became inactive because there was no treatment available.

Patients are typically thin and athletic, but there is a population of middle-aged women in there 40s and early 50s. They are symptomatic with sitting and ADLs and they benefit well as long as they don't have arthritis at the time of surgery. It was acknowledged that some patients are no doubt helped by this surgery, but there is no high level evidence of its effectiveness.

Craft invited public comment but there was none provided.

Action: A motion was made and seconded to approve the draft coverage guidance as appears below. **Motion approved 5-0 (Absent: Keenen).**

This version will now go out for public comment. Shaffer noted Herzka had submitted a number of articles which Dr. Little will review. These will be considered at the same time the public comments are reviewed.

HERC COVERAGE GUIDANCE

Surgery for femoroacetabular impingement syndrome is not recommended for coverage (*strong recommendation*).

B. Upper Endoscopy for GERD

Shaffer introduced the evidence included in the packet. He noted this involves very common symptoms and while endoscopy gives a more definitive diagnosis, the more scopings performed the more expense there is. One of the basic concerns is where to draw the line in terms of not detecting malignancies by not doing endoscopies. He reviewed the box language proposed by staff which draws the line as age 50, but that is not based strictly on the evidence. Endoscopy is recommended only after persistent symptoms following treatment under 50, but right away for those over 50. There is also a question about when to repeat endoscopies, which is suggested to be at a 9-year interval.

Shaffer said it is difficult to find the benefit from early endoscopy when looking through the studies. Overall there does not appear to be a clinically relevant benefit for UI for GERD symptoms. There are malignancy risks, which some studies reported to increase for males >35 and females >57. Authors thought age 55 was the most logical, but the subcommittee can discuss what age that should be. As for repeat endoscopy, if the initial test is negative, there was insufficient evidence to find benefit within the term of the study, which was 9 years. Endoscopy is relatively safe, but other treatment strategies are less costly. The only other consideration is whether to require a failure of PPI therapy, test and treat protocol for H pylori, or both in the younger patient.

Shaffer indicated he did get an email response from Drew Schembre, MD, the expert appointed on this topic who is expected to call in. Schembre basically agreed with the draft recommendations, however he did have concerns about patients on long-term PPI who would go without endoscopy if their symptoms are managed, especially those who are higher risk of esophageal cancer (e.g., smokers, family history of Barrett's, familial polyposis syndrome).

Mackay thought "uninvestigated dyspepsia" is a broad term and would like to see it changed to "chronic dyspepsia." "Recurrent" or "clinically significant" or a reference to alarm symptoms was also suggested as alternative language. It was agreed to add a disclaimer that this guidance does not apply to coverage of endoscopy for patients presenting with alarm symptoms and list the examples of anemia, weight loss and dysphasia (hemotenesis was also considered but dismissed).

Waldmann suggested a reference to "recurrent" or "chronic" symptoms, while MacKay proposed the language "unless the patient has persistent symptoms following completion of an appropriate course...."

At this time Schembre came on by phone. Craft indicated the subcommittee was leaning towards not having two categories of age ranges and instead requiring everyone to have a trial of PPI unless they have classic alarm symptoms. Schembre cautioned that this may be stepping into dangerous waters. The risk for asymptomatic precancerous development is significantly higher in older group, especially white males. It is those whose symptoms are well controlled who you should be worried about. If it was just a matter of treating the reflux you wouldn't look at all, but you are also trying to identify people at greatest risk for cancer and other problems who should be followed in surveillance program. A trial of PPI helps select out those who do well and the focus can turn to looking for indicators to select those at highest risk.

Discussion returned to the interval of repeat endoscopy. Nine years seemed arbitrary to some, with five years maybe making some sense. Others questioned whether to repeat the test at all. If someone has an endoscopy and the worry is that they have Barrett's, they never need another one unless their symptoms change.

Waldmann asked whether chewing tobacco would be in a high risk category and Schembre thought it probably would. Some also question whether those smoking marijuana in modern Washington would also fit, but it has not been studied.

It was noted that half who have Barrett's have no symptoms and are not taking PPIs. There is also a significant subgroup taking OTC medications who have low-grade symptoms and don't make much of it. The total prevalence of Barrett's varies but there is a 2% prevalence among asymptomatic white males. The notion of symptoms driving detection of Barrett's is limited. It is debated, but rate of Barrett's progressing to cancer is probably about 0.2 to 0.3 percent per year (12,000 new cases of adenocarcinoma per year).

Schembre said almost all esophageal cancers start as Barrett's, as they have a 40x higher risk. A small subgroup may not have Barrett's. There are well defined risk factors. The incidence of esophageal cancer is rising in all populations, for a variety of reasons, including H. pylori, increasing body masses and other environmental factors, but there is still a 4 or 5 to 1 male-to-female ratio. If you're looking at the topic of identifying unneeded upper endoscopies, a 30-year old professional under a lot of stress has virtually zero risk. Many would say let's treat but others would say let's do an endoscopy first and then talk, which is what we are trying to prevent. Someone with longstanding reflux deserves endoscopy and acid suppression; otherwise you would never have an opportunity to identify people with Barrett's who have been on OTC medications for 10 years. Schembre offer to provide literature on the evidence behind scoping that population as well as broader references on incidence, prevalence, trends and risk factors.

Craft asked the subcommittee what they would like to do. It could be sent out for comment based on the evidence; the expert comments don't have to be incorporated first. Waldmann made a case for separating the ages out and others agreed. Craft found a rate of 0.5 percent with Barrett's get cancer.

Little was asked where the 9-year interval came from, to which she responded a cohort study that doesn't really address the question. Ahmann has always heard 5-10 years, while Craft pointed out ASGE says no need to repeat test if initial one is negative. ASGE also says to repeat it every 3 years for those with Barrett's, while Ahmann has failed to see one convert to cancer over his 30 years of practice, yet everyone gets nervous about it.

While the evidence is lacking, it was agreed to leave the interval at years with the idea it can be altered based on public comment.

Action: A motion was made and seconded to approve the draft guidance as follows to be released for public comment. **Motion approved 5-0 (Absent: Keenen).**

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 completion of an appropriate course of 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).

Repeat endoscopy within nine years is not recommended for coverage for patients with dyspepsia after non-malignant findings on initial endoscopy (weak recommendation).

This guidance does not apply to coverage of endoscopy for patients presenting with “alarm symptoms” including, but not limited to, anemia, weight loss, and dysphagia.

C. Use of DXA in screening for and monitoring of osteoporosis

Shaffer felt this was a complex topic. Again, the more screening we do the more expense there is. The initial draft guidance is an attempt at trying to come up with some reasonable limits. The core source for this draft guidance is a USPSTF review published in 2010 and a NICE guideline. We were looking for evidence on frequency for those with both normal and low bone density. The additional source of Gourley et al was used to address frequency. DXA is the gold standard for diagnosing osteoporosis and fracture prediction.

No RCTs show the effect of screening on the fracture rate or the morbidity or mortality associated with fractures. However, evidence from drug studies have been reviewed in developing the evidence summary. Biphosphates (and others) have been shown to be effective in primary prevention, implying that if you can screen and get the appropriate individuals started on treatment, you can reduce the risk of fracture. The evidence is better for vertebral fractures and better in women than men. There is little risk in the use of DXA; rather the main risk is related to the drug treatment that follows. USPSTF says risk is relatively small compared to the risks of osteoporosis. There was inadequate evidence in men for use in primary prevention. The NICE guideline, which the fourth paragraph derives from, identifies risk categories that don't come from the evidence summary.

While the USPSTF identifies categories for increased use of DXA for measuring bone density, it does not address frequency. The Gourley study prospectively followed 4950 women over 66, categorized by initial bone density, and came up with rates of progression to osteoporosis which was used as the basis for the frequencies shown in the draft guidance. Shaffer then reviewed grade table and pointed out the Policy Landscape section that included the American College of Rheumatology recommendation that screening not be performed more often than every 2 years, even for those at high risk, and an interval as long as 10 years for healthy women >67 with normal bone density.

Waldmann noted the fourth paragraph addresses current or recent corticosteroids, but is silent on chronic use of steroids. If that were to be added what would be the dose? And what about inhaled steroids for asthma and allergic rhinitis? Muday pulled up the FRAX tool. It list glucocorticoids as a major risk factor – current oral use or previous exposure for at least 3

months at a dose equivalent to 5 mg or more and no mention of inhaled use. Waldmann thought this might be a good addition.

Michael Lewiecki, MD, the appointed expert on this topic said 10 million Americans have osteoporosis and another 4 million have osteopenia. Fractures have serious consequences: pain, disability, death and fracture related healthcare costs, despite the fact that we have excellent drugs. The problem is that most patients with osteoporosis are not being diagnosed, many of those diagnosed are not treated, and those prescribed are not taking the drugs long enough. It doesn't make much sense to him to create administrative roadblocks to a test that is underutilized.

Muday feels that concern that people who are appropriate for the screening aren't getting screened should not mean we should restrict over-screening. We see this in breast cancer. People who get screened get many screens. What we would like to see is the screens more widely used and not repeated so frequently.

Lewiecki addressed the question of who ought to be screened. If the initial screen shows good bone density and the individual has a low fracture risk, then you can wait a long time. But you need to screen appropriately in the first place. If, however, the fracture risk is high, you need to monitor the effects of treatment to make sure you are getting the desirable effect. Most will follow-up in one year and continue monitoring until stable and improved. Ahmann felt most wait for 2 years before retesting. Lewiecki said that while some do, he would not want to take medicine for 2 years without feedback about whether it is working or not.

At this time Shaffer reviewed the process, explaining that Dr. Lewiecki's comments will be added to those provided during the public comment period and the subcommittee will discuss in September as they work towards a final recommendation.

Lewiecki said the average Medicare reimbursement is \$50 per DXA. That's far below the cost of performing the test. Reimbursement is so low, many facilities are closing down and access is becoming difficult. It is not an expensive test so you may be penny wise and pound foolish to further restrict DXA.

Many members disagreed. We're making a big deal of it because we're finally getting the data to show outcomes. There is a wider margin of error than the changes in bone density it is supposedly detecting over short intervals. Also, if the patient is on a drug holiday the screening should be stopped. Even if the reimbursement is only \$50, that can add up because of the prevalence of the disease.

Ahmann asked whether there is evidence that men taking three months of prednisone gives them a higher risk of a fracture? Muday said it probably depends on other factors and it was agreed to remove the fourth paragraph of the draft presented. Furthermore, there was not felt to be the need to define "routine" screening in this case.

Keenen indicated DXA is sometimes used prior to surgery and it was clarified that this guidance instead applies to screening.

Action: A motion was made to approve and seconded to make changes to the initial draft coverage guidance resulting in the guidance below and post it for public comment.
Motion approved 6-0.

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 individuals with low bone mineral density, monitoring by repeat DXA scanning is not recommended for coverage more often than 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*).

5. PUBLIC COMMENT

No public comment was offered following the discussion of the three topics or at the end of the meeting.

6. ADJOURNMENT

The meeting was adjourned at 3:40 pm. The next meeting is scheduled for September 23, 2013 from 1:00-4:00 pm in Room 117B&C of the Meridian Park Hospital Community Health Education Center in Tualatin.

HEALTH EVIDENCE REVIEW COMMISSION (HERC)
COVERAGE GUIDANCE: HIP SURGERY PROCEDURES FOR
FEMOROACETABULAR IMPINGEMENT SYNDROME

DRAFT for HTAS Meeting Materials 9/23/2013

HERC COVERAGE GUIDANCE

Surgery for femoroacetabular impingement syndrome is not recommended for coverage (*strong recommendation*).

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

RATIONALE FOR GUIDANCE DEVELOPMENT

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

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

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

EVIDENCE SOURCES

Washington State Health Care Authority Health Technology Assessment Program. (2011). *Hip surgery procedures for treatment of femoroacetabular impingement syndrome: Health technology assessment*. Retrieved from <http://www.hta.hca.wa.gov/fai.html>

National Institute for Health and Clinical Excellence. (2011a). *Interventional Procedure Guidance 403: Open femoro-acetabular surgery for hip impingement syndrome*. London: National Institute for Health and Clinical Excellence. Retrieved from <http://guidance.nice.org.uk/IPG403>

3.0 FAI CG Draft-04-11-13

National Institute for Health and Clinical Excellence. (2011b). *Interventional Procedure Guidance 408: Arthroscopic femoro-acetabular surgery for hip impingement syndrome*. London: National Institute for Health and Clinical Excellence. Retrieved from <http://guidance.nice.org.uk/IPG408>

National Institute for Health and Clinical Excellence. (2011c). *Interventional procedure overview of open femoro-acetabular surgery for hip impingement syndrome*. London: National Institute for Health and Clinical Excellence. Retrieved from <http://guidance.nice.org.uk/IPG403>

National Institute for Health and Clinical Excellence. (2011d). *Interventional procedure overview of arthroscopic femoro-acetabular surgery for hip impingement syndrome*. London: National Institute for Health and Clinical Excellence. Retrieved from <http://guidance.nice.org.uk/IPG408>

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

SUMMARY OF EVIDENCE

Clinical Background

Femoroacetabular impingement (FAI) syndrome is a recently recognized diagnosis in primarily younger individuals where relatively minor abnormalities in the joint (orientation or morphology) are thought to cause friction/impingement and pain. It is theorized that FAI starts the breakdown of cartilage, leading to osteoarthritis. There are two types of FAI: cam impingement (non-spherical femoral head or abnormality at the head-neck junction) and pincer impingement (deep or retroverted acetabulum resulting in overcoverage of the femoral head). Proponents believe that surgical correction of the impinging deformities will alleviate the symptoms and retard the progression of osteoarthritis degeneration. Surgery to correct FAI includes arthroscopy, open dislocation of the hip and arthroscopy combined with a mini-open approach. The purpose of the surgery is to remove abnormal outgrowths of bone and damaged cartilage, and to reshape the femoral neck to ensure that there is sufficient clearance between the rim of the acetabulum and the neck of the femur. The causes of hip pain, the natural history of FAI and its relationship to osteoarthritis are unclear, and the case definition and selection criterion of patients for this procedure is uncertain. Furthermore, questions remain about the efficacy and effectiveness, safety and cost effectiveness of hip surgery for FAI.

Evidence Review

The evidence review addressed questions concerning case definition, evaluation of treatment outcomes, effectiveness and safety of hip surgery for FAI. To address the question of case definition, the most consistent case definition of FAI (cam or mixed) includes hip/groin pain, positive clinical impingement test, and an α -angle >50 - 55° . There is no evidence that the diagnosis of FAI can be obtained from clinical exam. One clinical test, the impingement sign, had a positive and negative predictive value of 86% and 79% in one study where the prevalence of FAI was 50%; however, in another study, the interobserver reliability of the impingement sign was only moderate. Even though the α -angle showed moderate to high interobserver reliability in several studies, it had poor diagnostic value in identifying FAI. Other imaging tests assessing abnormalities of the femur and acetabulum had variable degrees of reliability, but no others were tested for diagnostic validity.

Regarding outcome measures to evaluate the effectiveness of hip surgery for FAI, seven hip outcome measures were commonly used in the FAI patient population, but only three have undergone psychometric analysis in FAI (Hip Outcome Score, German version [HOS-D] and the modified Western Ontario and McMaster Universities Arthritis Index [M-WOMAC] or young hip-pain patients Nonarthritic Hip Score [NAHS]). Reliability was inadequately tested for all three instruments. The minimal clinically important difference was defined in only one measure, the HOS-D, and found to be 9 points for the activities of daily living subscale and 6 points for the sports subscale in FAI patients.

Regarding the efficacy of hip surgery for FAI, there are no randomized controlled trials (RCTs) available to assess the short- or long-term efficacy of FAI surgery compared with no surgery. Comparative evidence for this condition is limited to one retrospective cohort study comparing FAI surgery to conservative treatment, and four retrospective studies comparing various surgical treatments.

The only study to compare FAI surgery to no surgery included 17 patients (22 hips) who underwent three different treatments: nonoperative care with physiotherapy and anti-inflammatory medications, arthroscopy or open dislocation, and total hip replacement (THA). There were nine patients (10 hips) in the nonoperative group, six patients (eight hips) in the FAI surgery group, and two patients (four hips) in the THA group. The authors gave no indication of how these patients were selected or how many patients overall may have been eligible for the study; they simply stated that radiographic findings of osseous bump deformities on the anterolateral head-neck junction were found in all patients along with typical symptoms of FAI. They did, however, admit that the treatment received was based according to clinical and radiographic findings and MRI, thus acknowledging the potential of confounding by indication. Those with

moderate clinical symptoms but morphological signs of degenerative destruction of the hip joints underwent nonoperative treatment. Those with labral defects but only minor cartilage destruction on MRI underwent FAI surgery. The two patients who received THA did so as a result of having severe signs of osteoarthritis on radiographs. The authors provide no information regarding the patient selection process or loss to follow-up. There was no description of baseline characteristics apart from the mean age of patients. With respect to age, there were potentially important differences in ages of the patients among the three treatment groups. Only pain and return to work/sports are reported at final follow-up, with patients in the conservative group showing the poorest results overall: none were pain free at final follow-up compared with 100% of the patients in both surgical groups. Only 67% had returned to their previous work or sports level again compared with 100% of the patients in both surgical groups. It is difficult to draw any conclusions from this study as the patient groups compared were clearly different in many characteristics.

Of the other four cohort studies, two compared labral debridement with labral fixation, and two compared arthroscopic debridement and osteoplasty with arthroscopic debridement alone. Overall, none of these studies demonstrate that one specific treatment results in better outcomes than another (surgery versus no surgery, labral debridement versus refixation, osteoplasty versus no osteoplasty). Several case series report improvement in pain, patient reported and clinician reported hip outcome scores, patient satisfaction and return to normal activities following FAI surgery. However, whether this improvement is a result of the surgery, or the postoperative rehabilitation, or the change in activity subsequent to the surgery or placebo is not known. Approximately 8% of patients diagnosed with FAI who undergo surgery in published series go on to have a total hip arthroplasty within 3 years. There are no data available to assess long-term effectiveness of FAI surgery compared with no surgery. There are no data yet published to test the hypothesis that FAI surgery prevents or delays hip osteoarthritis or the need for total hip arthroplasty.

Regarding the safety of hip surgery for FAI, the risk of reoperation (other than conversion to THA) occurred in 4% (arthroscopy and open dislocation) and 9% of the patients (mini-open). There was only one reported head-neck fracture (0.1%) and no reports of AVN, osteonecrosis or trochanteric nonunion. Heterotopic ossification occurred in 2% to 3% of those receiving arthroscopy or mini-open, and 6% in those receiving open dislocation. Neurological complications (nerve palsy, paresthesia, and neuropraxia) were rare in those receiving arthroscopy or open dislocation; however, they occurred in 22% of 258 hips undergoing a mini-open procedure. Most were transient in nature.

[\[Evidence Source\]](#)

The National Institute for Clinical Excellence issued interventional procedure guidances on arthroscopic and open surgery for FAI in September and July 2011, respectively. Both guidances state that current evidence on the efficacy of arthroscopic or open femoro–acetabular surgery for hip impingement syndrome is adequate in terms of symptom relief in the short and medium term. With regard to safety, there are well recognized complications. Therefore this procedure may be used provided that normal arrangements are in place for clinical governance, consent and audit with local review of outcomes. They have established a registry to track long term outcomes of these procedures.

The literature review conducted to inform the NICE guidance consisted of one non-randomized controlled study and seven case series for the open procedure, and three non-randomized controlled studies, five case series and one case report for the arthroscopic procedure. The reviews report the following regarding the evidence base:

- Little or no controlled data are available comparing the procedure with other interventions or against natural history.
- A range of outcome assessment scales are used; validation of these scales is often not reported.
- The description of hip impingement pathology/lesions is not well defined in all studies.
- The intervention required is usually individualized to each patient, making comparison between studies difficult.

Study quality is generally poor, with little prospective data collection in case series.

[\[NICE IPG 403\]](#), [\[NICE IPG408\]](#)

Evidence Summary

The most consistent case definition of FAI (cam or mixed) includes hip/groin pain, positive clinical impingement test, and an α -angle >50 - 55° ; the predictive value of the impingement test ranges from 79 to 86%, while the α -angle has poor diagnostic value. Seven hip outcome measures are commonly used in the FAI patient population, but only three have undergone psychometric analysis in FAI, and reliability has been inadequately tested for all three. There are no data available to assess the short- or long-term efficacy of FAI surgery compared with no surgery, and no evidence that one specific treatment results in better outcomes than another. Regarding safety, the risk of reoperation (other than conversion to THA) is 4% to 9%, and heterotopic ossification occurs in 2% to 6% of patients, while neurological complications occur in up to 22% of patients.

After reviewing the available evidence including the lack of RCTs comparing FAI surgery to conservative care, as well as non-RCT comparative data demonstrating non-superiority of surgery, the WA HTA Clinical Committee concluded that the evidence was insufficient to recommend coverage of the procedure. The National Institute for Clinical Excellence has issued a guidance allowing for use of both arthroscopic and open procedures, despite a poor quality evidence base. They have established a registry to track long term outcomes.

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	Expert Input	Coverage Recommendation
Surgery for FAI	Unknown	Very low	Unknown at this time. Surgical intervention is generally more costly than conservative care in the short term.	High variability. Younger patients would generally want to avoid total hip replacement if there are effective alternatives.		Surgery for femoroacetabular impingement syndrome is not recommended for coverage (<i>strong recommendation</i>)

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

Note: GRADE framework elements are described in Appendix A

POLICY LANDSCAPE

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

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.

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.

Appendix B. Applicable Codes

CODES	DESCRIPTION
ICD-9 Diagnosis Codes	
715-715.9	Osteoarthritis
718.05	Articular cartilage disorder, pelvic region
718.45	Contracture of joint, pelvic region and thigh
718.65	Unspecified intrapelvic protrusion acetabulum, pelvic region and thigh
718.85	Other joint derangement, not elsewhere classified
718.95	Unspecified derangement of joint
719.45	Pain in joint, pelvic region and thigh
719.55	Stiffness of joint, not elsewhere classified, pelvic region and thigh
719.7	Difficulty in walking
719.85	Other specified disorders of joint, pelvic region and thigh
719.95	Unspecified disorder of joint, pelvic region and thigh
736.30	Acquired deformities of hip, unspecified deformity
736.39	Acquired deformities of hip, other
ICD-9 Volume 3 (Procedure Codes)	
80.15	Other Arthrotomy, Hip
80.25	Arthroscopy, Hip
80.45	Division Of Joint Capsule, Ligament, Or Cartilage; Hip
81.40	Repair Of Hip, Not Elsewhere Classified
CPT Codes	
29914	Arthroscopy, hip, surgical; with femoroplasty (i.e., treatment of cam lesion)
29915	Arthroscopy, hip, surgical; with acetabuloplasty (i.e., treatment of pincer lesion)
29916	Arthroscopy, hip, surgical; with labral repair
HCPCS Level II Codes	
None	

Note: Inclusion on this list does not guarantee coverage

Appendix C. HERC Guidance Development Framework

Surgery for FAI Syndrome

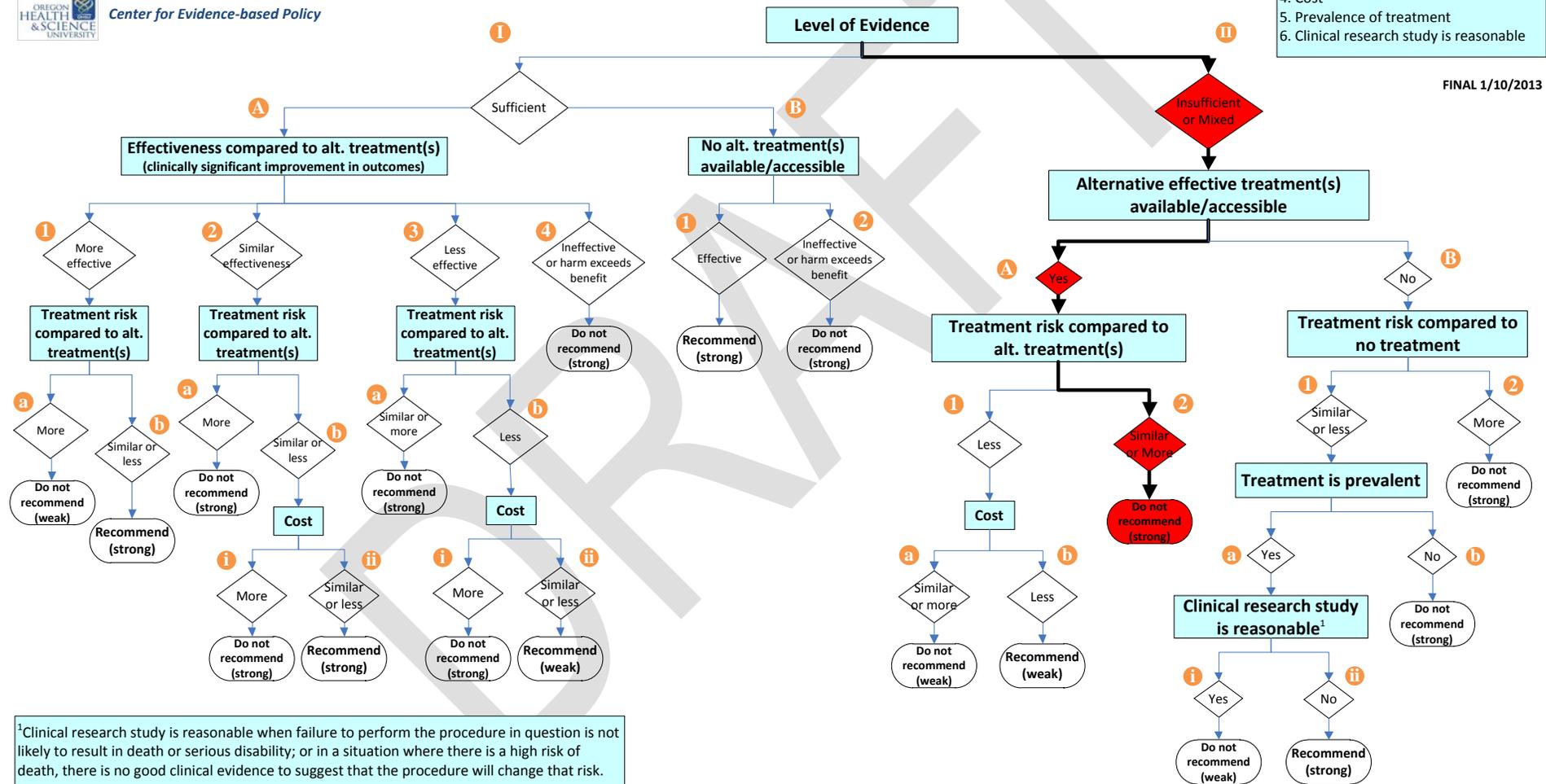


HERC Guidance Development Framework

Refer to HERC Guidance Development Framework Principles for additional considerations

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

FINAL 1/10/2013



HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Table of Contents

Commenters..... 1

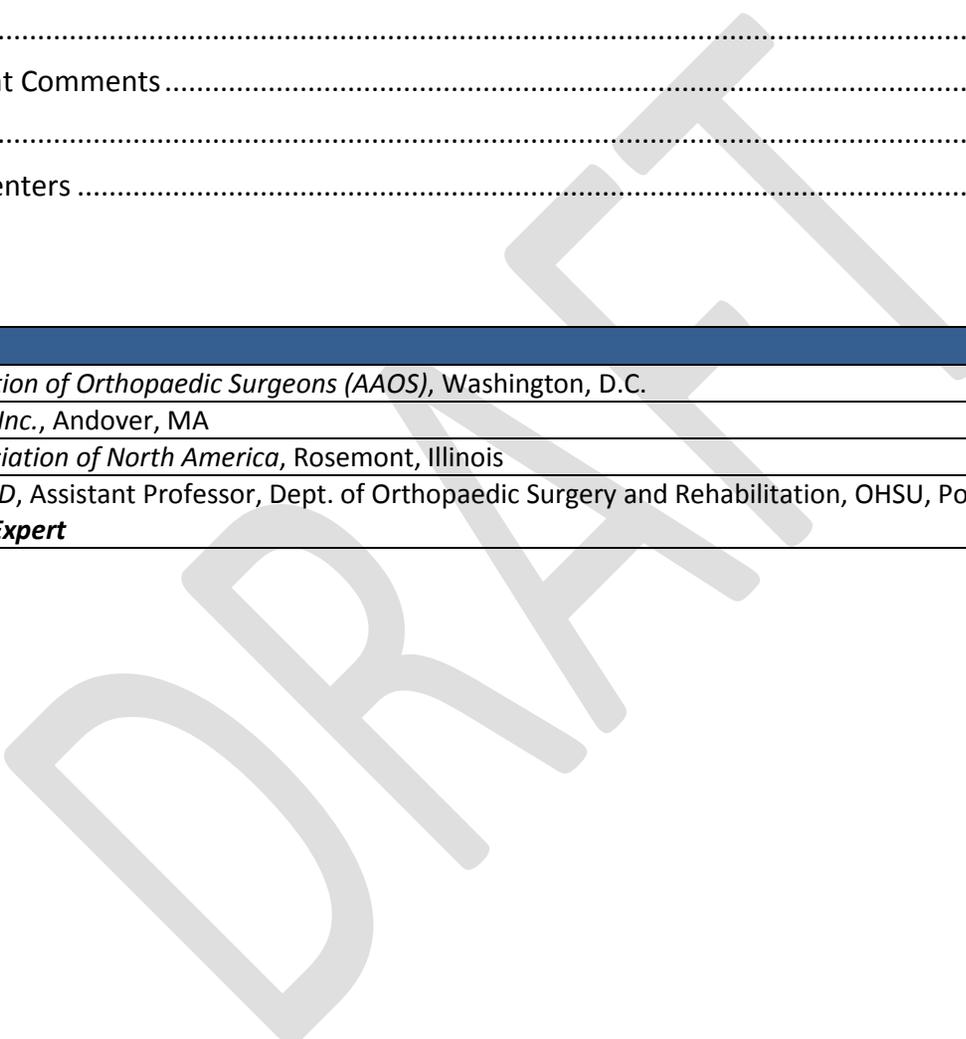
Discussion Questions & Relevant Comments..... 2

Public Comments 3

References Provided by Commenters 16

Commenters

Identification	Stakeholder
A	<i>American Association of Orthopaedic Surgeons (AAOS), Washington, D.C.</i>
B	<i>Smith & Nephew, Inc., Andover, MA</i>
C	<i>Arthroscopy Association of North America, Rosemont, Illinois</i>
D	<i>Andrea Herzka, MD, Assistant Professor, Dept. of Orthopaedic Surgery and Rehabilitation, OHSU, Portland, OR</i> HERC Appointed Expert



HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Discussion Questions & Relevant Comments

1. Is the very low quality evidence sufficient to recommend coverage?

NICE and Washington HTA differed on the sufficiency of the evidence, with NICE finding it weak but sufficient for coverage with a registry. Many commenters believe case series evidence to be sufficient. See comments ([4](#), [5](#), [15](#), [21](#), [22](#), [23](#), [28](#))

2. Are alternative effective treatments available/accessible for persons with FAI?

Expert comment indicates physical therapy is sufficient for some (but not all) patients with FAI-related hip pain, but no evidence was reviewed regarding the effectiveness of physical therapy or other nonsurgical alternatives for those who do not benefit from physical therapy. See comments ([14](#), [24](#), [32](#), [36](#), [37](#))

3. Is conducting a higher-quality study reasonable despite prevalence of treatment?

Expert testimony is that a randomized trial is not reasonable due to prevalence of the treatment. See comments ([25](#), [34](#), [35](#), [36](#))

4. What is the risk compared to alternative/no treatment?

Commenters describe risk of significant disability for patients who fail to benefit from physical therapy and do not have FAI surgery. However, surgery includes some risk of complications. See comments ([9](#), [25](#), [33](#), [34](#))

5. How do cost, along with patient values and preferences, affect the subcommittee's decision after reviewing the evidence?

Commenters cite the disabling nature of this condition, and its effects on younger patients which could result in high costs to society and patient preference for surgery. See comments ([17](#))

6. If the subcommittee chooses to recommend coverage, what are the appropriate indications for surgery? Should the subcommittee recommend coverage for resection and repair, or just repair?

See comments ([13](#), [24](#), [29](#))

7. Comments Designated for HTAS Discussion

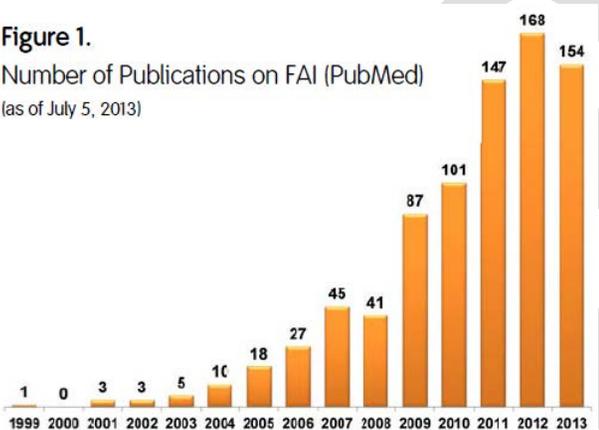
See comments ([6](#), [13](#), [15](#), [19](#), [21](#), [27](#), [30](#), [31](#), [32](#), [34](#), [37](#))

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Public Comments

Ident.	#	Comment	Disposition
A	1	<p>Thank you for the opportunity to comment on the draft guidance regarding hip surgery procedures for Femoroacetabular Impingement Syndrome (FAI). The American Association of Orthopaedic Surgeons (AAOS) represents 98% of the orthopaedic surgeons practicing in the United States, 368 of whom practice in Oregon. Orthopaedic surgeons are the preeminent physicians providing medical treatment of musculoskeletal conditions and disease.</p> <p>The AAOS firmly supports the incorporation of evidence into clinical practice, and is actively involved in developing and promoting Evidence Based Clinical Practice Guidelines for a number of musculoskeletal conditions. However, the AAOS opposes the proposed “no coverage” determination put forth by the Health Evidence Review Commission (HERC), because we do not believe this decision is consistent with evidence showing that surgery is a cost-effective treatment for the management of FAI Syndrome. Surgical treatment of FAI for symptomatic patients with ongoing disability issues can provide long- lasting symptom relief and allows these patients to return to work or other desired activities, reducing FAI’s economic burden on society.</p>	Thank you for this information.
A	2	The American Medical Association (AMA) concluded that FAI surgery is clinically effective; granting three Category 1 CPT codes effective January 2011. One criterion for granting Category 1 CPT codes is that “the clinical efficacy of the service/procedure is well established and documented in U.S. peer reviewed literature.” The AAOS believes that if a service or procedure has a Category I CPT code, it is not experimental or investigational. Therefore, payers should not deny reimbursement for these services and procedures when they are medically necessary. When payers do otherwise, they threaten the health of the public and unjustifiably interfere with the physician/patient relationship.	The existence of a Category I CPT code is not sufficient evidence of effectiveness.
A	3	All national U.S. commercial insurers and Medicare cover FAI surgery because it has been shown to be clinically effective. Since 2008, six independent systematic reviews of FAI surgery have concluded that published evidence supports its safety and effectiveness.	The HTAS is aware of this, but does not reach its conclusions based on the decisions of other payers. References not provided.
A	4	More than 40 peer-reviewed publications for symptomatic FAI using arthroscopic, open, or a combination of these surgical approaches report that patients’ symptoms are relieved and they are able to return to their normal activity levels.	References not provided. HTAS is unaware of any studies that were not included in the WA HTA report that are not case series. Case series are highly susceptible to bias and a lower quality type of evidence. Relates to discussion question #1
A	5	In July 2011, the National Institute for Health and Clinical Excellence in the United Kingdom stated in published guidance on open and arthroscopic FAI surgery, that current evidence of the efficacy of the procedures is adequate for the relief of associated symptoms.	HTAS is aware of the NICE guidance and it is included in the coverage guidance document. Their guidance acknowledges little or no controlled data comparing the procedure with other interventions or natural history. The structure of healthcare delivery in the UK allows them to

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition																																
			create a registry to track outcomes. HTAS does not have that ability and therefore needs to base its decisions on current evidence. <i>Relates to discussion question #1</i>																																
A	6	The AAOS once gain urges the Committee to revise its coverage guidance on hip surgery procedures for Femoroacetabular Impingement Syndrome (FAI) to be consistent with the other evidence-based coverage determinations and provide access to this safe, effective, and cost-effective treatment to Oregon’s public employees and Oregon Health Plan participants. Thank you for your consideration of these comments. Please do not hesitate to contact AAOS if we can be of further assistance.	<i>For HTAS discussion</i>																																
B	7	Smith & Nephew, Inc. is a global medical technology business specializing in Endoscopy, Orthopedics and Wound Management. We comment on the draft coverage guidance for surgery for Femoroacetabular Impingement (FAI) posted June 27, 2013. FAI understanding is evolving. Recognition of FAI as a disorder is a process in evolution. About 70 percent of all the literature on FAI has been published in 2010 or later. (Figure 1) Unrecognized, and/or inappropriately managed symptomatic FAI can lead to inefficient and wasteful use of medical resources. ¹ Figure 1. Number of Publications on FAI (PubMed) (as of July 5, 2013)  <table border="1" style="display: none;"> <caption>Data for Figure 1: Number of Publications on FAI (PubMed)</caption> <thead> <tr> <th>Year</th> <th>Number of Publications</th> </tr> </thead> <tbody> <tr><td>1999</td><td>1</td></tr> <tr><td>2000</td><td>0</td></tr> <tr><td>2001</td><td>3</td></tr> <tr><td>2002</td><td>3</td></tr> <tr><td>2003</td><td>5</td></tr> <tr><td>2004</td><td>10</td></tr> <tr><td>2005</td><td>18</td></tr> <tr><td>2006</td><td>27</td></tr> <tr><td>2007</td><td>45</td></tr> <tr><td>2008</td><td>41</td></tr> <tr><td>2009</td><td>87</td></tr> <tr><td>2010</td><td>101</td></tr> <tr><td>2011</td><td>147</td></tr> <tr><td>2012</td><td>168</td></tr> <tr><td>2013</td><td>154</td></tr> </tbody> </table>	Year	Number of Publications	1999	1	2000	0	2001	3	2002	3	2003	5	2004	10	2005	18	2006	27	2007	45	2008	41	2009	87	2010	101	2011	147	2012	168	2013	154	Thank you for taking the time to comment. HTAS is aware that a large volume of literature has been published since the date of the WA HTA report, but is unaware of any study type other than case series or retrospective cohort studies that were not included in that review (see comment #4).
Year	Number of Publications																																		
1999	1																																		
2000	0																																		
2001	3																																		
2002	3																																		
2003	5																																		
2004	10																																		
2005	18																																		
2006	27																																		
2007	45																																		
2008	41																																		
2009	87																																		
2010	101																																		
2011	147																																		
2012	168																																		
2013	154																																		
B	8	Pre-operative diagnosis. Contrary to the Health Technology Assessment published by Washington State, FAI can reasonably be diagnosed with high probability. ¹⁻⁴	<ul style="list-style-type: none"> Refs #1 and 2 were published before the date of the WA HTA report (last search date June 2011). The HTAS bases their guidance documents on reviews of the literature that utilize the highest standards of evidence 																																

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition
			<p>based medicine. Studies are included or excluded based on transparent, reproducible criteria; therefore the HTAS does not investigate individual studies. The HTAS assumes that the conclusions reached by the authors of these reviews weigh all the available evidence in accordance with the principles of evidence based medicine, and does not attempt to re-review the entire body of evidence to reach its own conclusions.</p> <ul style="list-style-type: none"> • Ref #3 is a SR of arthroscopic treatment of FAI. Authors report “We found that there was great inconsistency among the indications for arthroscopic management of FAI. Clinical and radiographic indices remain largely unvalidated.” • Ref #4 is a SR of treatment of FAI using open surgical dislocation. Authors state: “In short, there were major inconsistencies in the reported clinical and radiographic criteria used to indicate surgery among the 15 studies reviewed.” and “These results showed that that there was an inconsistency between the clinical and radiographic indications for surgical hip dislocation as a treatment for femoroacetabular impingement.” • Both of these reviews appear to contradict the commenter’s statement.
B	9	<p>Risk of delaying treatment for FAI.</p> <p>A recent evaluation of 561 consecutive hip arthroscopy patients (574 procedures: labral tear, 60.8%; FAI, 22.6%; condylar lesions, 16.6%) evaluated three patient segments by duration of symptoms: less than six months, six months to three years and over three years. Repeat arthroscopy on the same side or revision were more common in patients with delayed surgery⁵ (Figure 2).</p>	<p>This is a consecutive case series that compares outcomes based on length of symptoms. From abstract, unclear what kinds of baseline differences existed between groups, and whether they were controlled for.</p> <p><i>Relates to discussion question #4</i></p>

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition								
		<p>Figure 2. Hip Arthroscopy in 561 patients Relative risk of revision hip arthroscopy or arthroplasty on same side compared to patients with symptoms less than 6 months.</p> <table border="1" style="margin-left: auto; margin-right: auto;"> <caption>Data for Figure 2: Relative Risk of Revision Hip Arthroscopy or Arthroplasty</caption> <thead> <tr> <th>Symptom Duration</th> <th>Relative Risk</th> </tr> </thead> <tbody> <tr> <td>Under 6 months</td> <td>1</td> </tr> <tr> <td>6 months to 3 years</td> <td>2.75</td> </tr> <tr> <td>Over 3 years</td> <td>3.25</td> </tr> </tbody> </table>	Symptom Duration	Relative Risk	Under 6 months	1	6 months to 3 years	2.75	Over 3 years	3.25	
Symptom Duration	Relative Risk										
Under 6 months	1										
6 months to 3 years	2.75										
Over 3 years	3.25										
B	10	Philippon et.al. reported professional hockey players who delayed surgery beyond one year after acute injury were significantly slower in returning to sport. ⁶ Patients without access to joint preservation surgery who have unremitting symptomatic FAI may be left with total hip replacement as the only next step alternative.	The citation was published before the date of the WA HTA report (last search date June 2011) (see comment #8).								
B	11	Diagnostic recognition. The American Medical Association concluded FAI surgery was clinically effective and granted three Category Level 1 CPT codes effective January 2011.	See comment #2								
B	12	The United Kingdom’s National Institute for Health and Clinical Excellence (NICE) released guidance in September 2011 and July 2011, respectively, on arthroscopic and open surgery for FAI stating published evidence is adequate that surgery in symptomatic patients results in short- and medium-term benefits. ^{7,8}	See comment #5								
B	13	Health technology appraisals from all national commercial insurers recommend coverage in patients with symptoms and documented inability to participate in desired activities. Regence Blue Cross/Blue Shield covering Oregon, completed and published a Health Technology Assessment of FAI surgery in February 2013 and recommends it as “medically necessary to debride the bone when specific criteria are met.” ⁹	<p>No TA available on the TEC website. Citation is a medical coverage policy from BCBS, not a TA. Evidence review methods not specified and quality of review unknown. Evidence in the policy is summarized as follows:</p> <ul style="list-style-type: none"> • Not all patients with FAI morphology will have FAI pathology. • There is a high association between FAI pathology and idiopathic osteoarthritis, but this may represent a small proportion of the total cases of hip osteoarthritis. • Patients may present with hip pain that can be diagnosed as FAI by a combination of clinical evaluation, radiographs, and MR arthrography. 								

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome

Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition
			<ul style="list-style-type: none"> • In cases in which there is a positive impingement test result, anterosuperior labral or acetabular damage identified on MR arthrography and a pistol-grip morphology identified on imaging, there is a very high probability that the acetabular damage is caused by impingement of the femoral head-neck junction against the acetabular rim. FAI can be verified intraoperatively. • Repair of the labrum alone can improve symptoms in the short term. It is reasonable to expect that debridement/osteoplasty of the bump or bone spur would reduce continued abrasion in the long term. Some studies, albeit of low quality, support this view. • Treatment of FAI is most effective in younger patients without osteoarthritis (Tonnis grade 0 or I) or severe cartilage damage. Although osteoarthritis can be identified with plain film radiographs, articular damage is not always identified with current imaging techniques. • There is a high probability that symptoms in patients with osteoarthritis (Tonnis grade II or III, or joint space of less than 2 mm) or severe cartilage damage (Outerbridge grade IV) will not improve following osteoplasty. These patients may require THA for progressing pain within 5 years. • In large case series, arthroscopic treatment of FAI in young to middle-age patients without osteoarthritis and showing mild to moderate cartilage damage results in 75% to 85% of patients improved. • Smaller case series suggest that open treatment of FAI in young to middle-age patients with moderate to severe cartilage damage results in 50% to 70% of patients improved. Non-union has been reported to occur in 27% of patients following the transection of the great trochanter with hip dislocation.

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition
			<p>The literature is uncertain with respect to the following:</p> <ul style="list-style-type: none"> • It is not known whether arthroscopic or open approaches result in better net health outcomes when patients are matched for severity of FAI morphology and articular cartilage damage. • It is not known which patients with FAI morphology are most likely to progress to osteoarthritis. The progression of pincer impingement with damage initially restricted to the labrum may follow a different time course than cam-type impingement. • It is not known whether treatment of FAI will reduce the occurrence of osteoarthritis. <p><i>For HTAS discussion</i> <i>Relates to discussion question #6</i></p>
B	14	<p>Conservative management is ineffective.</p> <p>A just published systematic review reports, “Outcomes of operative treatment of femoroacetabular impingement are significantly better than nonsurgical management.”¹⁰ Non-surgical treatment of symptomatic FAI does not provide permanent symptom relief, may require permanent lifestyle modification and fails to allow patients to return to desired activity levels.^{6,11-28}</p>	<ul style="list-style-type: none"> • Ref #10 (Harris 2013) is a SR that includes 29 studies, overall quality score was poor. All study types with a minimum 2 year FU were eligible for inclusion. 83% were case series, total N=2369. While the author reports statistics to support the superiority of operative over non-operative treatment, there was only one study of non-operative management, which was a case series with n=37. No direct comparative evidence is reported. • Ref #6, 11-25 and 27 were published before the date of the WA HTA review (see comment #8). • Ref #26 is a case series of 23 patients evaluating the use of a specific hip distractor in the OR rather than a traction table. • Ref #28 is a retrospective case series of 47 high level athletes who underwent arthroscopic treatment of FAI. There was a 30% loss to follow up. The evaluable patients had significant improvements in pain (17/100 points) and function (12/100 points) scores (generally accepted minimum clinically important difference for HHS is 10).

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition
			<i>Relates to discussion question #2</i>
B	15	<p>Surgery relieves symptoms and allows patients to return to activity.</p> <p>Over 46 peer-reviewed publications for symptomatic FAI using arthroscopic, open or a combination of these surgeries report patients’ symptoms are relieved and the majority of patients are capable of returning to their previous level of activity.^{6;11-13;15-21;23;26-62} <u><i>There are no unfavorable reports.</i></u> Arthroscopic surgery for FAI was associated with the lowest overall risk of complications.</p> <p>Among these publications, 21 reports with collectively over 1300 patients document favorable surgical outcomes in 75 to 100 percent of symptomatic FAI patients who had failed non-surgical management comprised of medication, reduced activity and physical therapy or rehabilitation programs lasting up to and over one-year. Typical patients have been able to return to recreational and work activities within months and professional athletes have had their careers extended.^{6;11-15;17-23;26-28;45;53;55;58;61}</p>	<p>Minimum Clinically Important Difference (MCID) has not been defined according to the WA HTA report for either mHHS or NASH.</p> <ul style="list-style-type: none"> • Refs #6-23, 27, 29-52 and 56 were published before the date of the WA HTA review (see comment #8). • For Ref #26 and 28, see comment #14. • Ref #53 is another case series of 200 athletes, median improvement in MHHS was 20 points. • Ref #54 is a retrospective cohort study comparing labral repair with labral resection, and reported more improvement in those undergoing labral repair. • Ref #55 is a case series of 100 treated arthroscopically, median improvement was 21 points. • Ref #57 is a case series of 120 treated with minimally invasive anterolateral approach. Mean improvement in non-arthritis hip score (NASH) was 32 points. • Ref #58 is a case series of 44 athletic patients treated with the mini-open approach. Mean HHS improved 24 points. • Ref #59 is case series of 185/233 hips with 5 year FU treated with surgical hip dislocation. 82% were satisfied with the surgery. 14 hips underwent THA or major revision. • Ref #60 is SR that includes 31 studies of “generally low methodologic quality” and concludes “arthroscopy, open surgery and arthroscopic surgery followed by mini open surgery are comparable for functional results, biomechanics, and return to sport. Debridement and osteoplasty provide better results than debridement only. Significantly improved outcomes have been recorded in patients undergoing labral refixation than resection.” • Ref #61 is case series of 60 patients < age 17. Mean

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition
			<p>improvement in MHHS was 34. 13% required a second surgery for adhesions.</p> <ul style="list-style-type: none"> Ref #62 is a case series of 153 patients > age 50. 20% required THA (time period not clear). <p>For HTAS discussion Relates to discussion question #1</p>
B	16	<p>Since 2008, seven independent systematic reviews of FAI surgery for symptomatic patients each concluded that published evidence support its safety and effectiveness.^{10;63-68} Additional favorable reports have subsequently been published.^{17;28;53;54;57-62}</p>	<ul style="list-style-type: none"> For Ref #10, see comment #14. Ref #17 and 63-68 were published before the date of the WA HTA report (see comment #8). For Refs #28, 53, 54 and 57-62, see comment #15.
B	17	<p>FAI surgery is reported to be cost-effective.</p> <p>In a cost-effectiveness analysis⁶⁹ based upon best available data for patients with symptomatic FAI, observation or arthroscopic repair followed by hip replacement is compared to an endpoint of delaying total hip replacement surgery. FAI surgery was determined to be very cost-effective by the definition of cost-effectiveness used by the World Health Organization. It was more likely to demonstrate value in patients with limited pre-existing osteoarthritis or if progression to end-stage osteoarthritis is delayed.</p>	<p>Ref #69 is a CEA that reports an ICER of \$21,700 assuming a 3 year benefit of arthroscopy and no impact of treatment on natural history (no delay of progression to OA).</p> <p>Relates to discussion question #5</p>
B	18	<p>FAI surgery is right for patients with unremitting symptomatic FAI.</p> <p>FAI may be asymptomatic in many patients and all patients with FAI may not progress to osteoarthritis in the short-term. Additionally, outcomes are less favorable in the presence of preexisting advanced osteoarthritis of the affected hip.^{6;11;20;29;30;46;70;71}</p>	<p>All referenced citations were published before the date of the WA HTA report (see comment #8).</p>
B	19	<p>HERC solicited orthopedic expert Dr. Andrea Herzka to review the evidence. She stated on the Expert Review Form submitted April 22, 2013, "...parameters for medical necessity must be established to allow this population to receive the current standard of care for FAI which is arthroscopic intervention."</p> <p>The strong recommendation your guidance proposes to not cover FAI surgery is contrary to the conclusions from the best available evidence, your expert reviewer, Medicare and commercial insurers. Failure to cover hip surgeries for FAI will prevent patients who are suffering from chronic pain and disability from access to a surgery unanimously found reasonable, safe, effective and medically necessary. Patients with unremitting pain from symptomatic FAI and no hip preservation surgery option may ultimately seek hip replacement, a more costly alternative. We urge you to act in the best interest of your patients and cover FAI surgery in symptomatic patients meeting appropriate criteria.</p>	<p>HTAS is aware of Dr. Herzka's opinion, as she has provided both oral and written testimony to the committee on two occasions.</p> <p>HTAS believes that this is a new diagnosis that previously has been untreated or treated by other means. Given the inadequate evidence base to assess efficacy and harms compared to conservative treatment, the HTAS has elected a non-coverage decision with the hope that good quality research will be conducted to better inform their policy decision soon.</p> <p>For HTAS discussion</p>
C	20	<p>The Arthroscopy Association of North America (AANA) is an Accredited Council for Continuing</p>	<p>Thank you for your comment and your interest in</p>

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition
		Medical Education approved organization which exists to promote, encourage, support and foster, through continuing medical education functions, the development and dissemination of knowledge in the discipline of surgery. This is done to improve upon the diagnosis and treatment of diseases and injuries of the musculoskeletal system to enhance the lives our patients. AANA has 3691 members across the United States, Canada and Mexico dedicated to this mission. We are recognized internationally as leaders in teaching the management of musculoskeletal disease states and have been using evidence based methodologies to support that teaching for over 30 years. We welcome the recent trend toward more evidence based practice that the Oregon Health Authority has adopted. AANA works closely with the American Academy of Orthopedic Surgeons in formulating Clinical Practice Guidelines and Appropriate Use Criteria to assist orthopedic surgeons in providing treatment recommendation that incorporate the best available evidence in the medical literature.	evidence based practice.
C	21	<p>The recent recommendation published for comment by the Health Technology Assessment Subcommittee (HTAS) of the Oregon State Health Evidence Review Commission (HERC) on the surgical treatment of femoro-acetabular impingement (FAI) troubles us. We believe it violates the tenets of evidence based medicine and the currently recognized standard of care for patients with symptomatic FAI in the Oregon medical community as well as nationally.</p> <p>Evidence based medicine seeks to improve patient care by 1) using clinical expertise, 2) searching the scientific literature for the best available studies to evaluate and compare treatments, and 3) include patient values and preferences in recommending therapies.¹ The HERC has severely limited the inclusion of all the medical evidence that supports the surgical treatment of FAI, discounted expert clinical opinion, and not included at all the aspect of patient preferences in its strong recommendation against coverage for these procedures.</p>	<p>Ref #1 also states “It is when asking questions about therapy that we should try to avoid the non-experimental approaches, since these routinely lead to false positive conclusions about efficacy. Because the randomised trial, and especially the systematic review of several randomised trials, is so much more likely to inform us and so much less likely to mislead us, it has become the "gold standard" for judging whether a treatment does more good than harm.”</p> <p><i>For HTAS discussion</i> <i>Relates to discussion question #1</i></p>
C	22	There are only two evidence sources cited for the recommendation; the Washington State Healthcare Authority Technology Assessment ² and the National Institute for Health and Clinical Excellence (NICE) ³ both from 2011. The WSCH HTA severely circumscribed evidence synthesis rules to favor randomized clinical trials (RCT) as demonstration of procedure efficacy. This methodology ignores the over 40 peer-reviewed case-controlled and case series publications demonstrating improved outcomes from FAI surgery in symptomatic patients that this young, active population of patients experiences when conservative care fails. By stressing RCTs as the only measure of clinical efficacy, the HERC does not allow for a true assessment of the surgical literature as it exists today, limiting patient access to important therapies.	<p>The WA HTA report is a full systematic review. It is unclear what the commenter means by “WSCH HTA severely circumscribed evidence synthesis rules to favor randomized clinical trials (RCT) as demonstration of procedure efficacy.” The tenets of evidence-based medicine are clear that, in general, RCTs are required to draw valid conclusions, because to do otherwise results in significant propensity for bias. Indeed, nearly all Cochrane reviews limit their inclusion criteria to RCTs.</p> <p><i>Relates to discussion question #1</i></p>
C	23	The NICE report actually states, “Current evidence on the efficacy of femoro–acetabular surgery for hip impingement syndrome is adequate in terms of symptom relief in the short and medium term.” ³	<p>See comment #5</p> <p><i>Relates to discussion question #1</i></p>

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition
		The NICE recommendation cited does not justify a strong recommendation against covering these procedures.	
C	24	A recent high quality review of the surgical treatment of FAI not included in the HTAS review states, “Outcomes of operative treatment of femoroacetabular impingement are significantly better than nonsurgical management. Surgical treatment significantly improves outcomes, with no consistent significant differences exhibited between open and arthroscopic techniques. Open surgical dislocation has significantly greater reoperation and complication rates vs. mini-open and arthroscopic techniques. Outcomes of labral refixation are significantly better than debridement in patients with labral injuries.” ⁴	This is the same SR (Harris 2013) previously addressed in comment #14. Relates to discussion question #2 Relates to discussion question #6
C	25	As recognized experts in the treatment of these conditions, the members of AANA consider the surgical treatment of symptomatic FAI to be <i>the standard of care</i> ⁵ for those patients that fail the conservative management of these conditions. The HTAS recommendation against coverage violates this standard and places these patients at unnecessary risk of further pain, disability and potentially irreversible joint destruction. ^{6,7} Since FAI surgery is recognized as a standard, there is no clinical equipoise ⁸ to support conducting RCTs. Indeed, “If the clinician knows, or has good reason to believe, that a new therapy (A) is better than another therapy (B), he cannot participate in a comparative trial of Therapy A versus Therapy B. Ethically, the clinician is obligated to give Therapy A to each new patient with a need for one of these therapies.” ⁹	The history of medicine is full of examples of therapies that were thought to be better than another, yet were proven not to be by a RCT. Examples include stem cell transplant for breast cancer, hormone replacement therapy for menopausal women and many drugs that have been removed from the market because of unacceptable risk. Many people have been harmed by these treatments that were approved/ performed without RCTs. Relates to discussion question #3 Relates to discussion question #4
C	26	Lastly, Medicare and private insurers do cover surgery for FAI ¹⁰ . To deny coverage to patients for these procedures who are covered by Oregon State creates a potential treatment disparity for the poor and minority patients served by state programs such as Medicaid.	HTAS is aware of the coverage policies of other payers (see comment #3).
C	27	Evidence based methodologies are necessary to help improve patient care and make treatment more consistent with the current state of medical knowledge. It is important to have experts examine guidelines to offer necessary insight concerning their relevance and veracity. The members of AANA hope that you will reconsider your coverage decision to avoid the standard of care and treatment disparity issues we have described. We would be happy to advise the HERC on further guidelines concerning musculoskeletal healthcare to improve the care that all the citizens of Oregon deserve.	For HTAS discussion
D	28	The proposed draft for coverage for FAI is outdated and flawed. The literature used to create this draft comprised of articles published prior to 2011. More recent citations that I provided to the committee have not yet been incorporated into this guideline. Although the level of evidence of many outcome studies is level IV, their cumulative value is powerful. The safety and efficacy of hip arthroscopy for FAI associated non arthritic hip pain is well established.	None of the evidence previously submitted provides direct comparison between operative and non-operative management. Relates to discussion question #1

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition
D	29	A specific example of the inaccuracy of the proposed draft is the statement that there is no evidence to support one surgical treatment over another when Dr. Larson, Dr. Philippon and Dr. Ganz have both published that labral preservation results in significantly improved outcomes. As the expert on FAI for the state of Oregon, I cited several articles relevant to this guideline and published since the Washington State Health Care Authority Health Technology Assessment Program recommendation in early 2011.	Several cohort studies were provided previously that suggest differences in efficacy based on surgical technique, as reported in the Harris SR (see comment #14). The WA HTA report concludes there is no evidence that one specific treatment, including surgical approach, results in better outcomes. In addition, this does not address the efficacy of surgery compared to non-operative management.
D	30	I provided a powerpoint presentation and strongly encouraged the subcommittee to create a category for medical necessity that would allow for coverage for this disabled patient population after failed conservative treatment and when all other criteria for medical necessity have been met. This is an effective way to allow disabled patients to have access to newer technologies when there are no known alternative treatment options.	<i>For HTAS discussion</i> <i>Relates to discussion question #6</i>
D	31	The authors also state that proponents of surgery for FAI “believe that surgical correction of the impinging deformities will alleviate the symptoms and retard the progression of osteoarthritis degeneration.” This is incorrect. The current literature supports surgical intervention for patients with symptomatic FAI refractory to conservative treatment. The surgical goal is pain reduction and improved function. These outcomes measures have been successfully reproduced and documented in short and midterm outcome studies. The safety and efficacy of surgical intervention for FAI is very well documented. In patients with refractory non-arthritic hip pain due to FAI related chondrolabral care surgical intervention is the current standard of care. We are hopeful that long term data will <i>also</i> demonstrate a change in the natural history and retardation of arthritic progression, but at this time this data remains unknown. For this reason, FAI surgery is not indicated as a prophylactic surgery in individuals without significant disability and pain. The uncertainty of the natural history or the impact of surgery on arthritic progression is irrelevant to this guidance recommendation. This outcome measure can not be known for several years, but the efficacy of pain reduction and improved function and level of activity are outcome measures that are well documented in short and mid term outcome studies.	Thank you for this clarification regarding the current indications for surgical intervention. <i>For HTAS discussion</i>
D	32	The algorithm used by the commission is incorrect. Under “alternative effective treatments available” the committee chooses “yes.” I am unaware of any alternative effective treatment options for FAI related non arthritic hip pain in patients refractory to conservative treatment. My patients have all failed a combination of physical therapy, massage, steroid injections, acupuncture, chiropractic care, heat, ice, NSAIDS, and at times opiates prior to surgical intervention.	Thank you for providing your opinion on this matter. <i>For HTAS discussion</i> <i>Relates to discussion question #2</i>

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition
D	33	Nonarthritic hip pain associated with FAI can be debilitating. Many young patients are unable to walk, sit or perform their job duties. When conservative treatment fails in this patient population, surgical intervention becomes a matter of medical necessity. BCBS, aetna, Tricare, healthnet, lifewise, MODA, pacificsource, cigna, and providence insurance carriers have all updated their coverage guidelines for FAI based on the literature to cover surgery when “medical necessity” criteria are met. Although these criteria differ very slightly from one another, they are conceptually identical. The Washington State Health Care Authority Health Technology Assessment Program is the only group of reviewers to determine a non-coverage guideline and this was based on literature through the latter part of 2010. It is the only review to date that denies its patients of the current standard of care for the treatment of FAI.	HTAS is aware of the coverage policies of other payers (see comment #3). Relates to discussion question #4
D	34	The data for FAI supports the following progression through the current HERC algorithm. Alternative treatment avail?- NO Treatment risk compared to no treatment- similar or less (difficult to quantify chronic pain and loss of function, weight gain due to inactivity and the cost of unemployment, obesity, depression and chronic pain) but the minor known risks of FAI surgery seem to outweigh the consequences of untreated pain Treatment is prevalent- YES – covered by all commercial insurance carriers and performed in every European country, Canada, Mexico, and Brazil, Korea, and Japan. Clinical research study is reasonable- NO	<i>For HTAS discussion</i> Relates to discussion question #3 Relates to discussion question #4
D	35	Patients who have already failed months of PT and activity modification should not be subjected to further lack of intervention. The orthopedic surgery societies are all in agreement that denial of surgical intervention after failed conservative treatment is not ethical given the profound disability that these patients suffer.	A RCT similar to those completed for viscosupplementation of the knee, and lavage and debridement of the knee for OA would seem reasonable; in the latter example, lavage and debridement was compared with sham surgery. HERC has made coverage recommendations in both of these procedures. In both of those examples, patients likely had exhausted prior treatment options with the exception of joint replacement, similar to the situation with FAI. Relates to discussion question #3
D	36	Certainly research to better understand the efficacy of conservative treatment of FAI is warranted. Those patients who have not trialed any conservative therapy should be studied, but unfortunately, this population might not be well represented in the orthopedic literature because the majority of patients who present to my office have already failed a course of PT and are desperate for pain relief. The primary care and/or physical therapy providers will likely need to conduct this type of study since	This comment suggests that FAI can be treated conservatively with PT. Organized systems of care may be more likely to be able to accomplish this, such as the Kaiser system or the VA. Relates to discussion question #2

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

Ident.	#	Comment	Disposition
		<p>they are likely treating patients with FAI associated pathology successfully with PT. Those patients rarely make it to my office and their exclusion creates a bias. In addition, patients successfully treated with PT often cancel follow up since they are improved.</p> <p>Finding a primary care practice with a high enough volume of FAI patients with hip pain to study prospectively is a huge hurdle to research. Capturing this patient population is extremely difficult and we will never know how many patients with mild pain self treat successfully with activity modification.</p>	<p><i>Relates to discussion question #3</i></p>
D	37	<p>Prospective evaluation of conservative treatment of FAI is challenging for the reasons explained above. For this reason, I have recommended that the subcommittee uses the algorithm to make a strong “yes” recommendation for the surgical treatment of FAI associated non arthritic hip pain in patients with symptoms refractory to conservative treatment.</p> <p>Those few patients whom I see in my office who have not trialed a formal course of PT with avoidance of flexion and focus on strengthening are treated initially with this conservative care. If treatment fails after 3-6 months, then surgery is recommended. It is my experience that a minority of patients are successfully managed without surgery, and that increased activity level/athleticism and younger age are associated with high failure rate of conservative care.</p> <p>Lastly, the subcommittee wanted to know how patients with this condition have been managed historically prior to this intervention. Many patients had to live with “a bum wheel.” They could no longer go for a walk with friends, sit or stand at their jobs, or have sex without severe pain. These patients were often told “there was nothing wrong” because their X-rays did not demonstrate arthritis and nobody could explain this elusive groin pain. These patients lived with chronic pain and fear of hip rotation and flexion. Without an explanation for their pain they likely suffered both physically and emotionally.</p>	<p>Thank you for providing your opinion on this matter.</p> <p><i>For HTAS discussion</i></p> <p><i>Relates to discussion question #2</i></p>
D	38	<p>Dr. Ganz, the first surgeon to describe FAI came to this concept after performing hundreds of hip replacements in young patients with premature advanced arthritis and years of preceding pain with similarly misshapen hips. This finding guided his philosophy that this abnormal morphology caused painful injuries to the cartilage and labrum in the hip joint and led to eventual arthritis: thus, the introduction of the concept of FAI.</p>	<p>Thank you for this information.</p>

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

References Provided by Commenters

A	None submitted
B	<p>References</p> <ol style="list-style-type: none"> 1. Clohisy JC, Knaus ER, Hunt DM, Leshner JM, Harris-Hayes M, Prather H. Clinical presentation of patients with symptomatic anterior hip impingement. <i>Clin Orthop Relat Res</i> 2009;467:638-644. 2. Philippon MJ, Maxwell RB, Johnston TL, Schenker M, Briggs KK. Clinical presentation of Femoroacetabular impingement PHILIPPON2007A. <i>Knee Surg Sports Traumatol Arthrosc</i> 2007;15:1041-1047. 3. Ayeni OR, Wong I, Chien T, Musahl V, Kelly BT, Bhandari M. Surgical Indications for Arthroscopic Management of Femoroacetabular Impingement. <i>Arthroscopy</i> 2012. 4. Ayeni OR, Naudie D, Crouch S et al. Surgical indications for treatment for femoroacetabular impingement with surgical hip dislocation AYENI2012. <i>Knee Surg Sports Traumatol Arthrosc</i> 2012. 5. Aprato A, Jayasekera N, Villar R. Timing in hip arthroscopy: does surgical timing change clinical results? <i>Int Orthop</i> 2012;36:2231-2234. 6. Philippon MJ, Weiss DR, Kuppersmith DA, Briggs KK, Hay CJ. Arthroscopic labral repair and treatment of Femoroacetabular impingement in professional hockey players. <i>Am J Sports Med</i> 2010;38:99-104. 7. NICE. Open femoro-acetabular surgery for hip impingement syndrome. IPG403. 6-15-2012. National Institute for Health and Clinical Excellence. 7-5-2013. 8. NICE. Arthroscopic femoro-acetabular surgery for hip impingement syndrome. [IPG408]. 6-15-2012. National Institute for Health and Clinical Excellence. 7-5-2013. 9. Regence Blue Cross & Blue Shield. Femoroacetabular Impingement Surgery. <i>Regence Medical Policy Manual</i> 2013;Policy No. 160:1-17. 10. Harris JD, Erickson BJ, Bush-Joseph CA, Nho SJ. Treatment of femoroacetabular impingement: a systematic review. <i>Curr Rev Musculoskelet Med</i> 2013. 11. Jager M, Wild A, Westhoff B, Krauspe R. Femoroacetabular impingement caused by a femoral osseous head-neck bump deformity: clinical, radiological, and experimental results. <i>J Orthop Sci</i> 2004;9:256-263. 12. Espinosa N, Rothenfluh DA, Beck M, Ganz R, Leunig M. Treatment of femoro-acetabular impingement: preliminary results of labral refixation. <i>J Bone Joint Surg Am</i> 2006;88:925-935. 13. Bizzini M, Notzli HP, Maffiuletti NA. Femoroacetabular impingement in professional ice hockey players: a case series of 5 athletes after open surgical decompression of the hip. <i>Am J Sports Med</i> 2007;35:1955-1959. 14. Heyworth BE, Shindle MK, Voos JE, Rudzki JR, Kelly BT. Radiologic and intraoperative findings in revision hip arthroscopy. <i>Arthroscopy</i> 2007;23:1295-1302. 15. Philippon M, Schenker M, Briggs K, Kuppersmith D. Femoroacetabular impingement in 45 professional athletes: associated pathologies and return to sport following arthroscopic decompression. <i>Knee Surg Sports Traumatol Arthrosc</i> 2007;15:908-914. 16. Stahelin L, Stahelin T, Jolles BM, Herzog RF. Arthroscopic offset restoration in femoroacetabular cam impingement: accuracy and early clinical outcome. <i>Arthroscopy</i> 2008;24:51-57. 17. Byrd JW, Jones KS. Arthroscopic femoroplasty in the management of cam-type femoroacetabular impingement. <i>Clin Orthop Relat Res</i> 2009;467:739-746. 18. Graves ML, Mast JW. Femoroacetabular impingement: do outcomes reliably improve with surgical dislocations? <i>Clin Orthop Relat Res</i> 2009;467:717-723. 19. Yun HH, Shon WY, Yun JY. Treatment of Femoroacetabular impingement with surgical dislocation. <i>Clin Orthop Surg</i> 2009;1:146-154. 20. Gedouin JE, May O, Bonin N et al. Assessment of arthroscopic management of femoroacetabular impingement. A prospective multicenter study. <i>Orthop Traumatol Surg Res</i> 2010;96:S59-S67. 21. Haviv B, O'Donnell J. Arthroscopic treatment for symptomatic bilateral cam-type femoroacetabular impingement. <i>Orthopedics</i> 2010;33:874. 22. Meermans G, Konan S, Haddad FS, Witt JD. Prevalence of acetabular cartilage lesions and labral tears in femoroacetabular impingement. <i>Acta Orthop Belg</i> 2010;76:181-188.

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

23. Singh PJ, O'Donnell JM. The outcome of hip arthroscopy in Australian football league players: a review of 27 hips. *Arthroscopy* 2010;26:743-749.
24. Byrd JW. Arthroscopic hip surgery for the treatment of femoroacetabular impingement. *Orthopedics* 2011;34:186.
25. Emara K, Samir W, Motasem EL, Ghafar KA. Conservative treatment for mild femoroacetabular impingement. *J Orthop Surg (Hong Kong)* 2011;19:41-45.
26. Flecher X, Dumas J, Argenson JN. Is a hip distractor useful in the arthroscopic treatment of femoroacetabular impingement? *Orthop Traumatol Surg Res* 2011;97:381-388.
27. Javed A, O'Donnell JM. Arthroscopic femoral osteochondroplasty for cam femoroacetabular impingement in patients over 60 years of age. *J Bone Joint Surg Br* 2011;93:326-331.
28. Nho SJ, Magennis EM, Singh CK, Kelly BT. Outcomes after the arthroscopic treatment of femoroacetabular impingement in a mixed group of high-level athletes. *Am J Sports Med* 2011;39Suppl:145-195.
29. Beck M, Leunig M, Parvizi J, Boutier V, Wyss D, Ganz R. Anterior femoroacetabular impingement: part II. Midterm results of surgical treatment. *Clin Orthop Relat Res* 2004;67-73.
30. Murphy S, Tannast M, Kim YJ, Buly R, Millis MB. Debridement of the adult hip for femoroacetabular impingement: indications and preliminary clinical results. *Clin Orthop Relat Res* 2004;178-181.
31. Peters CL, Erickson JA. Treatment of femoro-acetabular impingement with surgical dislocation and debridement in young adults. *J Bone Joint Surg Am* 2006;88:1735-1741.
32. Beaulé PE, Le Duff MJ, Zaragoza E. Quality of life following femoral head-neck osteochondroplasty for Femoroacetabular impingement. *J Bone Joint Surg Am* 2007;89:773-779.
33. Peters CL, Schabel K, Anderson L, Erickson J. Open treatment of femoroacetabular impingement is associated with clinical improvement and low complication rate at short-term followup. *Clin Orthop Relat Res* 2010;468:504-510.
34. Naal FD, Miozzari HH, Wyss TF, Notzli HP. Surgical hip dislocation for the treatment of Femoroacetabular impingement in high-level athletes. *Am J Sports Med* 2011;39:544-550.
35. Endo H, Noda T, Mitani S et al. Operative treatment for pincer type femoroacetabular impingement: a case report. *Acta Med Okayama* 2010;64:149-154.
36. Hartmann A, Gunther KP. Arthroscopically assisted anterior decompression for femoroacetabular impingement: technique and early clinical results. *Arch Orthop Trauma Surg* 2009;129:1001-1009.
37. Laude F, Sariali E, Nogier A. Femoroacetabular impingement treatment using arthroscopy and anterior approach. *Clin Orthop Relat Res* 2009;467:747-752.
38. Lincoln M, Johnston K, Muldoon M, Santore R. Combined arthroscopic and modified open approach for cam femoroacetabular impingement: a preliminary experience. *Arthroscopy* 2009;25:392-399.
39. Nepple JJ, Zebala LP, Clohisy JC. Labral Disease Associated With Femoroacetabular Impingement: Do We Need to Correct the Structural Deformity? *J Arthroplasty* 2009;24:114-119.
40. Pierannunzii L, d'Imporzano M. Treatment of femoroacetabular impingement: a modified resection osteoplasty technique through an anterior approach. *Orthopedics* 2007;30:96-102.
41. Ilizaliturri VM, Jr., Nossa-Barrera JM, Acosta-Rodriguez E, Camacho-Galindo J. Arthroscopic treatment of femoroacetabular impingement secondary to paediatric hip disorders. *J Bone Joint Surg Br* 2007;89:1025-1030.
42. Bardakos NV, Vasconcelos JC, Villar RN. Early outcome of hip arthroscopy for femoroacetabular impingement: the role of femoral osteoplasty in symptomatic improvement. *J Bone Joint Surg Br* 2008;90:1570-1575.
43. Ilizaliturri VM, Jr., Orozco-Rodriguez L, Acosta-Rodriguez E, Camacho-Galindo J. Arthroscopic treatment of cam-type femoroacetabular impingement: preliminary report at 2 years minimum follow-up. *J Arthroplasty* 2008;23:226-234.

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

44. Larson CM, Giveans MR. Arthroscopic management of femoroacetabular impingement: early outcomes measures. *Arthroscopy* 2008;24:540-546.
45. Philippon MJ, Yen YM, Briggs KK, Kuppersmith DA, Maxwell RB. Early outcomes after hip arthroscopy for Femoroacetabular impingement in the athletic adolescent patient: a preliminary report. *J Pediatr Orthop* 2008;28:705-710.
46. Sampson TG. Arthroscopic treatment of Femoroacetabular impingement. *Am J Orthop (Belle Mead NJ)* 2008;37:608-612.
47. Brunner A, Horisberger M, Herzog RF. Sports and recreation activity of patients with femoroacetabular impingement before and after arthroscopic osteoplasty. *Am J Sports Med* 2009;37:917-922.
48. Larson CM, Giveans MR. Arthroscopic debridement versus refixation of the acetabular labrum associated with femoroacetabular impingement. *Arthroscopy* 2009;25:369-376.
49. Philippon MJ, Briggs KK, Yen YM, Kuppersmith DA. Outcomes following hip arthroscopy for femoroacetabular impingement with associated chondrolabral dysfunction: minimum two-year follow-up. *J Bone Joint Surg Br* 2009;91:16-23.
50. Gedouin JE, Duperron D, Langlais F, Thomazeau H. Update to femoroacetabular impingement arthroscopic management. *Orthop Traumatol Surg Res* 2010;96:222-227.
51. Horisberger M, Brunner A, Herzog RF. Arthroscopic treatment of femoroacetabular impingement of the hip: a new technique to access the joint. *Clin Orthop Relat Res* 2010;468:182-190.
52. Sionek A, Czubak J, Polaczek P, Czwojdzinski A. Osteochondroplasty as a treatment of Femoroacetabular impingement. *Ortop Traumatol Rehabil* 2010;12:504-510.
53. Byrd JW, Jones KS. Arthroscopic management of femoroacetabular impingement in athletes. *Am J Sports Med* 2011;39 Suppl:7S-13S.
54. Schilders E, Dimitrakopoulou A, Bismil Q, Marchant P, Cooke C. Arthroscopic treatment of labral tears in Femoroacetabular impingement: a comparative study of refixation and resection with a minimum two-year follow-up. *J Bone Joint Surg Br* 2011;93:1027-1032.
55. Byrd JW, Jones KS. Arthroscopic Management of Femoroacetabular Impingement: Minimum 2-Year Follow-up. *Arthroscopy* 2011;27:1379-1388.
56. Sampson TG. Arthroscopic Treatment of Femoroacetabular Impingement. *Techniques in Orthopaedics* 20[1], 56-62. 2005.
57. Chiron P, Espie A, Reina N, Cavaignac E, Molinier F, Laffosse JM. Surgery for femoroacetabular impingement using a minimally invasive anterolateral approach: analysis of 118 cases at 2.2-year follow-up. *Orthop Traumatol Surg Res* 2012;98:30-38.
58. Cohen SB, Huang R, Ciccotti MG, Dodson CC, Parvizi J. Treatment of Femoroacetabular Impingement in Athletes Using a Mini-Direct Anterior Approach. *Am J Sports Med* 2012.
59. Naal FD, Miozzari HH, Schar M, Hesper T, Notzli HP. Midterm Results of Surgical Hip Dislocation for the Treatment of Femoroacetabular Impingement. *Am J Sports Med* 2012.
60. Papalia R, Del BA, Franceschi F, Marinozzi A, Maffulli N, Denaro V. Femoroacetabular impingement syndrome management: arthroscopy or open surgery? *Int Orthop* 2012;36:903-914.
61. Philippon MJ, Ejnisman L, Ellis HB, Briggs KK. Outcomes 2 to 5 Years Following Hip Arthroscopy for Femoroacetabular Impingement in the Patient Aged 11 to 16 Years. *Arthroscopy* 2012.
62. Philippon MJ, Schroder E Souza BG, Briggs KK. Hip arthroscopy for femoroacetabular impingement in patients aged 50 years or older. *Arthroscopy* 2012;28:59-65.
63. Bedi A, Chen N, Robertson W, Kelly BT. The management of labral tears and femoroacetabular impingement of the hip in the young, active patient. *Arthroscopy* 2008;24:1135-1145.
64. Clohisy JC, St John LC, Schutz AL. Surgical treatment of femoroacetabular impingement: a systematic review of the literature. *Clin Orthop Relat Res* 2010;468:555-564.
65. Ng VY, Arora N, Best TM, Pan X, Ellis TJ. Efficacy of surgery for femoroacetabular impingement: a systematic review. *Am J Sports Med* 2010;38:2337-2345.
66. Stevens MS, Legay DA, Glazebrook MA, Amirault D. The evidence for hip arthroscopy: grading the current indications. *Arthroscopy* 2010;26:1370-1383.
67. Botser IB, Smith TW, Jr., Nasser R, Domb BG. Open surgical dislocation versus arthroscopy for Femoroacetabular impingement: a comparison of clinical outcomes.

HERC Coverage Guidance – Hip Surgery Procedures for Femoroacetabular Impingement (FAI) Syndrome Disposition of 2nd Round of Public Comments

	<p><i>Arthroscopy</i> 2011;27:270-278.</p> <p>68. Matsuda DK, Carlisle JC, Arthurs SC, Wierks CH, Philippon MJ. Comparative systematic review of the open dislocation, mini-open, and arthroscopic surgeries for Femoroacetabular impingement. <i>Arthroscopy</i> 2011;27:252-269.</p> <p>69. Shearer DW, Kramer J, Bozic KJ, Feeley BT. Is Hip Arthroscopy Cost-effective for Femoroacetabular Impingement? <i>Clin Orthop Relat Res</i> 2011.</p> <p>70. Horisberger M, Brunner A, Herzog RF. Arthroscopic treatment of femoral acetabular impingement in patients with preoperative generalized degenerative changes. <i>Arthroscopy</i> 2010;26:623-629.</p> <p>71. Kim KC, Hwang DS, Lee CH, Kwon ST. Influence of femoroacetabular impingement on results of hip arthroscopy in patients with early osteoarthritis. <i>Clin Orthop Relat Res</i> 2007;456:128-132.</p>
C	<p>References</p> <ol style="list-style-type: none"> 1. Sackett D. <i>Evidence-based medicine: What it is and what it isn't</i>. BMJ 1996; 312:71-77. 2. <i>Hip Surgery Procedures for Treatment of Femoroacetabular Impingement Syndrome</i>: Health Technology Assessment, August 26, 2011. Washington State Health Care Authority. WA Health Technology Assessment, Olympia, WA 98504-2712. 3. <i>IPG403 Open femoro-acetabular surgery for hip impingement syndrome: guidance</i>. July 27, 2011. National Institute for Health and Clinical Excellence, London. 4. Harris JD, Erickson BJ, Bush-Joseph CA, Nho SJ. <i>Treatment of femoroacetabular impingement: a systematic review</i>. <i>Curr Rev Musculoskelet Med</i> 2013 June 7. 5. <i>Position Statement on FAI</i>: Arthroscopy Association of North America, Rosemont IL. Shouldn't there be some sort of date on the position statement? 6. Aprato A, Jayasekera N, Villar R. <i>Timing in hip arthroscopy: does surgical timing change clinical results?</i> <i>Int Orthop</i> 2012;36:2231-2234. 7. Philippon MJ, Weiss DR, Kuppersmith DA, Briggs KK, Hay CJ. <i>Arthroscopic labral repair and treatment of femoroacetabular impingement in professional hockey players</i>. <i>Am J Sports Med</i> 2010; 38:99-104. 8. Freedman B. <i>Equipoise and the ethics of clinical research</i>. <i>N Engl J Med</i> 1987;317:141-145. 9. Shaw LW, Chalmers TC. <i>Ethics in Cooperative Clinical Trials</i>. <i>Ann N Y Acad Sci</i>, 1970;169(2):487-495. 10. <i>Position Statement on FAI</i>: Regence BCBS Oregon, Utah, Idaho, Washington, 2013.
D	None (references submitted previously)

HEALTH EVIDENCE REVIEW COMMISSION (HERC)

COVERAGE GUIDANCE: UPPER ENDOSCOPY FOR GASTROESOPHAGEAL REFLUX DISEASE (GERD) AND DYSPEPSIA SYMPTOMS

DRAFT for HTAS Meeting Materials 9/23/2013

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 endoscopy.](#)

[In the absence of significant new symptoms, r](#)Repeat endoscopy ~~within nine years~~ is not recommended for coverage for patients with dyspepsia [or GERD](#) after non-malignant findings on initial endoscopy (weak recommendation).

This guidance does not apply to coverage of endoscopy for patients presenting with “alarm symptoms” including, but not limited to, anemia, weight loss, and dysphagia.

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

RATIONALE FOR GUIDANCE DEVELOPMENT

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

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

Coverage guidance development follows to translate the evidence review to a policy decision. Coverage guidance may be based on an evidence-based guideline developed by the Evidence-based Guideline Subcommittee or a health technology assessment developed by the Health Technology Assessment Subcommittee. In addition, coverage

guidance may utilize an existing evidence report produced by one of HERC's trusted sources, generally within the last three years.

EVIDENCE SOURCE

Liu, R., Kriz, H., Thielke, A., Vandegriff, S., & King, V. (2012). *Upper endoscopy for gastroesophageal reflux disease (GERD) and upper gastrointestinal (GI) symptoms*. Olympia: Washington State Health Authority Health Technology Assessment Program. Retrieved February 21, 2013, from <http://www.hta.hca.wa.gov/gerd.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

Gastroesophageal reflux disease (GERD) is the most common outpatient gastrointestinal diagnosis in the United States, with a prevalence of 10% to 58.3% and an annual incidence of 0.38% to 0.45%. The Montreal consensus panel, an international Consensus Group tasked with developing a global definition and classification of GERD, reached strong consensus in defining GERD as "a condition which develops when the reflux of stomach contents causes troublesome symptoms and/or complications". Common symptoms of GERD include heartburn (defined as a burning sensation behind the breastbone), regurgitation and chest pain. Obesity; the presence of a hiatal hernia; and the use of estrogen, nitrates, anticholinergics, and tobacco products are considered risk factors for GERD. Gastroesophageal reflux disease can lead to a decreased quality of life and to more severe conditions such as esophagitis, Barrett's esophagus, and adenocarcinoma of the esophagus.

Dyspepsia is estimated to range in prevalence in the United States from 2.9 to 34.4%. The Rome III Committee defines dyspepsia as having one or more of the following symptoms: epigastric pain or burning; postprandial fullness; and/or early satiety. Other dyspeptic symptoms may include nausea and vomiting, upper abdominal bloating, heart burn, and regurgitation. Dyspepsia symptoms are distinguished from GERD as not being "troublesome" enough, referring to the Montreal definition of GERD; however, many authors have used the terms interchangeably.

The signs and symptoms of GERD, dyspepsia, and other more severe conditions such as Barrett's esophagus, can be very similar, and diagnostic procedures can be used to establish a diagnosis and rule out other possible conditions. Diagnostic procedures for dyspepsia and GERD can include questionnaires, empiric therapeutic trial, pH monitoring, upper endoscopy, and/or double contrast barium swallow. Empiric

Coverage Guidance: Upper Endoscopy for Gastroesophageal Reflux Disease (GERD) And Dyspepsia Symptoms

DRAFT for HTAS Meeting Materials 9/23/2013

4.0 UGI GERD Draft-9-11-13

therapeutic trial is a commonly employed strategy for patients presenting with GERD. This includes both an empiric trial of proton pump inhibitors (PPI) and test-and-treat for *Helicobacter pylori* (*H. pylori*). An empiric trial of PPIs typically includes twice daily dosing for four weeks, and a daily dose of 40 to 80mg of omeprazole is the most common PPI regimen used in clinical empiric therapy studies. The sensitivity and specificity of this PPI test ranges from 62 to 92% and 36 to 100%, respectively.

Evidence Review

Effectiveness of Early Treatment Strategies

With regard to the effectiveness of early treatment strategies for GERD that include upper endoscopy compared with empiric medical management, one good quality systematic review including two separate meta-analyses was identified. One evaluated early endoscopy versus empiric PPI and the other evaluated early endoscopy versus test-and-treat for *H. pylori*. The first meta-analysis included five RCTs and found no difference in symptomatic cure at 12 months between endoscopy and PPI arms. The second meta-analysis, also including five RCTs, was first done by pooling trial-level data. This analysis found no difference in effect (RR 0.95, 95% CI 0.79 to 1.15), but a high degree of statistical heterogeneity. When an alternate analysis of these same five studies was done using individual patient data, there was no longer statistical heterogeneity and a small but statistically significant benefit to upper endoscopy emerged (OR 0.75, 95% CI 0.58 to 0.96; RR 0.95, 95% CI 0.92 to 0.99).

A single fair quality prospective cohort study of 70 patients found that 24-hour pH monitoring is the most accurate single diagnostic test for GERD, when a concordance of three separate tests (omeprazole challenge, endoscopy, histology, pH monitoring) is taken as the gold standard. However, the authors note that there are barriers to its widespread use including invasiveness, cost, and availability. A serial application of an omeprazole challenge test, endoscopy, and finally histopathology achieves a sensitivity of 100% for GERD diagnosis.

Overall, considering all the available evidence from the systematic review plus the cohort study, there does not appear to be a clinically relevant benefit of prompt upper endoscopy over test and treat strategies or empiric PPI therapy for uninvestigated GERD symptoms in the primary care setting. (*Overall strength of evidence: High*)

Indications for Early Endoscopy

When considering whether there are clinical signs and symptoms that may be useful to identify patients for whom early endoscopy improves health outcomes, one good quality systematic review of 57,363 patients in 17 prospective cohort studies was identified.

They found that alarm symptoms¹, clinical opinion, and computer modeling programs based on symptom questionnaires were all unreliable predictors of gastrointestinal malignancy. Sensitivity ranged from 0% to 83% while specificity varied from 40% to 98%.

A good quality prospective cohort study found cancer in 0.9% of patients presenting with uncomplicated dyspepsia (i.e., without alarm symptoms) and the findings suggest that risk is correlated with age greater than 35 for males and greater than 57 for females. A fair quality prospective cohort study determined that American Society of Gastroenterologists (ASGE) guideline criteria (indications for endoscopy) were poorly correlated with clinically relevant endoscopic findings, although having a guideline indication does marginally increase the pre-test probability of endoscopy (from 45% to 47%), while not having one lowers it (from 45% to 29%). A second fair-quality prospective cohort study in the setting of open-access endoscopy found that 15% of the patients with esophagogastric carcinoma did not present with alarm symptoms and may have suffered delayed diagnosis without early endoscopy; however, there was an unusually high prevalence (3%) of cancer in the study population. Finally, a fair-quality prospective cohort study of primary care patients with uninvestigated dyspepsia found that Barrett's esophagus was most likely in patients who were male, greater than 50 years old, had symptoms of at least 5 to 10 years duration, and suffered predominantly from reflux.

The authors of the systematic review noted above suggest that, in the absence of compelling predictors, the concept of "alarm symptoms" should not be abandoned at this time. They suggest age greater than 55 as "the most logical alternative strategy... because the incidence of upper GI malignancy is negligible in Western populations at younger ages and only rises in prevalence above the age of 55 years." In contrast, the authors of the good quality cohort study suggest that age should be lower (35) for males and could be higher (57) for females. (*Overall strength of evidence: Moderate*)

Repeat Endoscopy

With regard to whether there are diagnoses for which repeat endoscopy is indicated, only one study, a prospective cohort study of good quality, addressed the question. This study evaluated the utility of repeat endoscopy in patients who initially presented with dyspeptic symptoms and had non-malignant endoscopic findings. About a third of these patients underwent a subsequent endoscopy within nine years of the index study. The results of these later endoscopies are not known; however, patients who had further endoscopy were neither more nor less likely than other patients to be symptomatic eight to nine years after the index study ($\chi^2=0.6$, $df=1$, $p > 0.05$). Overall, evidence is

¹ Weight loss, dysphagia, anemia

insufficient to suggest repeat endoscopy to any patients with initial dyspepsia who have non-malignant findings on their index endoscopy. (*Overall strength of evidence: Low*)

Harms of Endoscopy

None of the included studies addressed harms. According to the authors of one economic evaluation, most harms of endoscopy are cardiorespiratory in nature; that is, related to the procedure sedation rather than the endoscope itself. These authors used a 0.02% incidence of severe harms and modeled their economic assumptions on the surgical repair of perforation. No data was identified on harms associated with empiric acid-suppression or *H. pylori* test-and-treat. (*Overall strength of evidence: Insufficient*)

Subpopulations

Age was the only factor associated with differential effectiveness in one good quality meta-analysis. The authors of this study performed subgroup analyses based on age, gender, predominant symptom, and presence of *H. pylori*. There was a small but statistically significant benefit of endoscopy in patients 50 years of age and older (RR=0.90, 95% CI 0.82 to 1.00, $p < 0.05$); no other associations were found. A good-quality prospective cohort study found that patients with malignancy were on average 20 years older than patients without malignancy ($p < 0.001$). A fair-quality prospective cohort study also found increasing prevalence of malignancy with rising age. In a good quality economic evaluation simulation model, relative effectiveness of interventions was similar, but resulted in slightly fewer additional quality adjusted life years (QALY) in hypothetical 30 year olds than in hypothetical 60 year olds. A poor quality retrospective chart review of VA patients failed to find any correlation between significant endoscopic findings (Barrett's esophagus and/or erosive esophagitis) and age, gender, race, or NSAID use. (*Overall strength of evidence: Moderate [Age], Insufficient [All others]*)

Cost-effectiveness of Endoscopy Compared to Other Treatment Strategies

With the exception of empiric therapy for US 30 year olds, all five good quality studies, one of two fair quality studies, and one of three poor quality studies favored *H. pylori* test-and-treat as the most cost-effective strategy for adults with uninvestigated symptoms of dyspepsia and/or GERD. Only two studies, both of good quality, evaluated the cost-effectiveness of different management strategies for new upper gastrointestinal symptoms in a US population. In a simulation model, empiric PPI was the strategy of choice for 30 year old patients, and test-and-treat for *H. pylori* was the most cost-effective intervention for 60 year olds. A decision analysis looked only at patients less than 45 years of age, and determined that adding a 6-week trial of PPI to the test-and-treat strategy improved its cost-effectiveness. A good quality economic evaluation of Canadian individual patient data concluded that no one strategy was the most clearly cost-effective, but at a clinically relevant willingness-to-pay threshold of CAN\$30,000 to 70,000 per QALY, omeprazole treatment based on the CanDys protocol (which

Coverage Guidance: Upper Endoscopy for Gastroesophageal Reflux Disease (GERD) And Dyspepsia Symptoms

DRAFT for HTAS Meeting Materials 9/23/2013

4.0 UGI GERD Draft-9-11-13

incorporates test-and-treat for those without heartburn or reflux as the predominant symptom) was the most cost-effective. Two other good quality models also favored the test-and-treat approach, along with one fair and one poor quality RCT. One fair quality decision analysis favored a screening questionnaire followed by prompt endoscopy for high-risk patients. Two poor quality RCTs found empiric PPI to be the most cost-effective alternative, but did not include comparison to *H. pylori* testing and treatment. There were no economic studies that found prompt endoscopy to be the most cost-effective intervention. (*Overall strength of evidence: Moderate*)

[\[Evidence Source\]](#)

Evidence Summary

Overall, the evidence does not point to a clinically relevant benefit of prompt upper endoscopy over test-and-treat strategies or empiric PPI therapy for uninvestigated GERD symptoms in the primary care setting. Alarm symptoms, clinical opinion, and computer modeling programs based on symptom questionnaires are all unreliable predictors of gastrointestinal malignancy. The harms of endoscopy, or of any of the treatment strategies for GERD or dyspepsia, have not been well documented in this literature base. There is an increasing prevalence of malignancy with rising age, and there may be a small benefit of endoscopy over other initial treatment strategies in patients over 50 based on one trial. Test-and-treat for *H. pylori* is likely the most cost-effective strategy for adults with uninvestigated symptoms of dyspepsia and/or GERD.

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	Expert Input	Coverage Recommendation
Endoscopy as initial evaluation for new onset GERD or dyspepsia in patients under 50 or 55	No net benefit compared to other treatment strategies	High	Other treatment strategies are less costly	Low variability. Most would want to avoid endoscopy; some would prefer definite diagnosis before treatment	<i>Dr. Schembre agrees with the coverage recommendation, but has concerns regarding high risk individuals on long-term PPI therapy without endoscopy</i>	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 completed an appropriate course of PPI therapy or an H. pylori test and treat protocol (strong recommendation).
Endoscopy as initial evaluation for new onset GERD or dyspepsia in patients over 50 or 55	Small net benefit compared to other treatment strategies	Moderate	Endoscopy moderately more costly	Moderate variability	<i>Concurs with this recommendation</i>	Upper endoscopy for uninvestigated dyspepsia or GERD symptoms is recommended for coverage in patients at least 50 years of age (strong recommendation)

Indication	Balance between desirable and undesirable effects	Quality of evidence*	Resource Allocation	Values and preferences	Expert Input	Coverage Recommendation
Repeat endoscopy after initial endoscopy for GERD with non-malignant findings	No apparent benefit, small harms	Insufficient	Endoscopy moderately more costly	Moderate variability	<i>Does not support this recommendation. At the least, would recommend 5 years as a more reasonable restriction.</i>	Repeat endoscopy within nine years is not recommended for coverage for patients with dyspepsia after non-malignant findings on initial endoscopy (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

Eight quality measures were identified when searching the [National Quality Measures Clearinghouse](#). Five are sponsored by the American Gastroenterological Association Institute, while three are sponsored by the Australian Council on Healthcare Standards. None have been endorsed by the National Quality Forum. The five American measures are listed below:

1. Gastroesophageal reflux disease (GERD): percentage of patients aged 18 and older with the diagnosis of GERD who have been prescribed chronic proton pump inhibitor (PPI) or histamine H₂ receptor antagonist (H₂ RA) therapy who received an assessment of their GERD symptoms within 12 months
2. Gastroesophageal reflux disease (GERD): percentage of patients aged 18 seen for an initial evaluation of GERD who did not have a barium swallow test ordered
3. Gastroesophageal reflux disease (GERD): percentage of patients aged 18 and older with the diagnosis of GERD, seen for an initial evaluation, who were assessed for the presence or absence of the following alarm symptoms: involuntary weight loss, dysphagia, and GI bleeding
4. Gastroesophageal reflux disease (GERD): percentage of patients aged 18 and older with the diagnosis of GERD or heartburn whose endoscopy report indicates a suspicion of Barrett's esophagus who had a forceps biopsy performed
5. Gastroesophageal reflux disease (GERD): percentage of patients aged 18 and older seen for an initial evaluation with at least one alarm symptom who were either referred for upper endoscopy or had an upper endoscopy performed

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.

Coverage Guidance: Upper Endoscopy for Gastroesophageal Reflux Disease (GERD) And Dyspepsia Symptoms

DRAFT for HTAS Meeting Materials 9/23/2013

4.0 UGI GERD Draft-9-11-13

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.

Appendix B. Applicable Codes

CODES	DESCRIPTION
ICD-9 Diagnosis Codes	
530.1	Esophagitis
530.11	Reflux esophagitis
530.12	Acute esophagitis
530.19	Other esophagitis
530.2	Ulcer of esophagus
530.21	... with bleeding
530.3	Stricture and stenosis of esophagus
530.81	Esophageal reflux
530.85	Barrett's esophagus
530.89	Other specified disorders of esophagus
530.9	Unspecified disorder of esophagus
535	Gastritis and duodenitis
535.0	Acute gastritis
535.2	Gastric mucosal hypertrophy
535.3	Alcoholic gastritis
535.4	Other specified gastritis
535.5	Unspecified gastritis and gastroduodenitis
536.2	Persistent vomiting
536.8	Dyspepsia and other specified disorders of function of stomach
536.9	Unspecified functional disorder of stomach
786.5	Chest pain
786.59	Other chest pain
787.1	Heartburn
787.2	Dysphagia
787.21	...oral phase
787.22	... oropharyngeal phase
787.23	... pharyngeal phase
787.24	... pharyngoesophageal phase
787.29	Other dysphagia
789	Other symptoms involving abdomen and pelvis
789.06	... epigastric
789.07	... generalized
789.09	... other specified site
ICD-9 Volume 3 (Procedure Codes)	
42.23	Other esophagoscopy
42.24	Closed [endoscopic] biopsy of esophagus
44.13	Other gastroscopy

Coverage Guidance: Upper Endoscopy for Gastroesophageal Reflux Disease (GERD) And Dyspepsia Symptoms

DRAFT for HTAS Meeting Materials 9/23/2013

4.0 UGI GERD Draft-9-11-13

CODES	DESCRIPTION
44.14	Closed [endoscopic] biopsy of stomach
45.13	Other endoscopy of small intestine
45.14	Closed [endoscopic] biopsy of small intestine
45.16	Esophagogastroduodenoscopy [EGD] with closed biopsy
CPT Codes	
43200	Esophagoscopy, rigid or flexible; diagnostic, with or without collection of specimen(s) by brushing or washing (separate procedure)
43202	...with biopsy, single or multiple
43235	Upper gastrointestinal endoscopy including esophagus, stomach, and either the duodenum and/or jejunum as appropriate; diagnostic, with or without collection of specimen(s) by brushing or washing (separate procedure)
43239	...with biopsy, single or multiple
HCPCS Level II Codes	
None	

Note: Inclusion on this list does not guarantee coverage

Endoscopy for Evaluation of Dyspepsia/GERD after age 50/55 (Compared to PPI or Test and Treat)

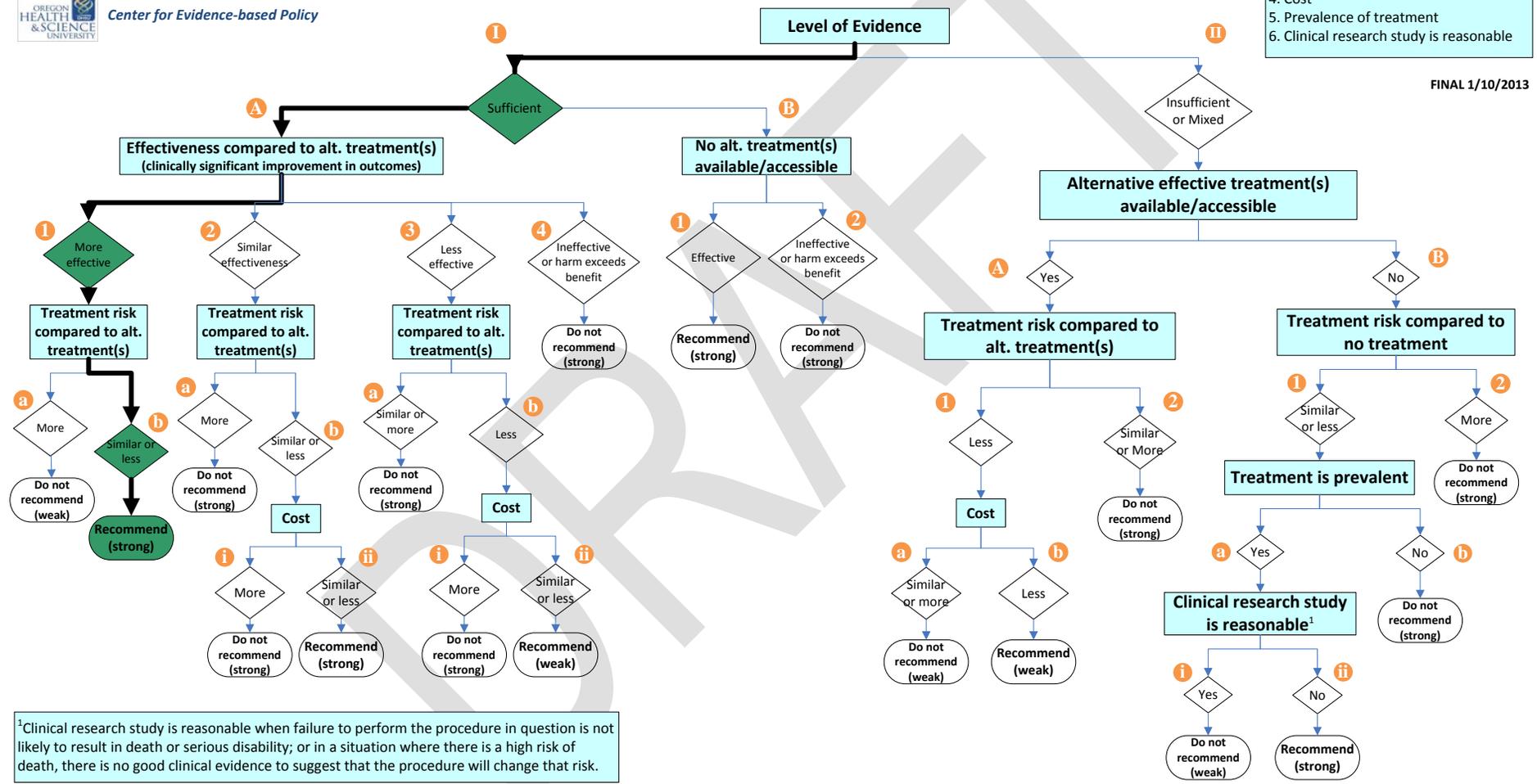


HERC Guidance Development Framework

Refer to HERC Guidance Development Framework Principles for additional considerations

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

FINAL 1/10/2013



Repeat endoscopy within 9 years after initial endoscopy for GERD with non-malignant findings

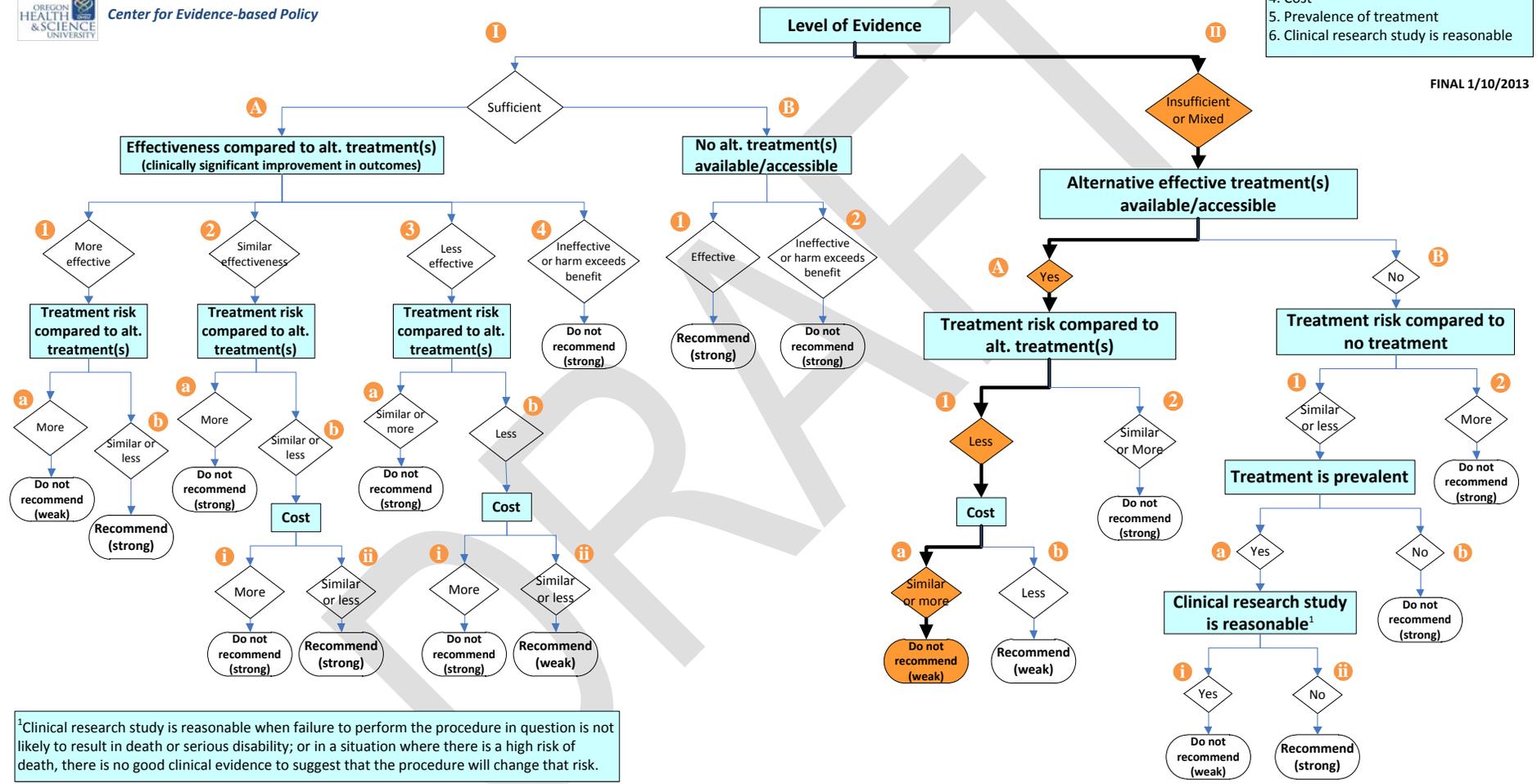


HERC Guidance Development Framework

Refer to HERC Guidance Development Framework Principles for additional considerations

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

FINAL 1/10/2013



HERC Coverage Guidance – Upper Endoscopy for Gastroesophageal Reflux Disease (GERD) and Dyspepsia Symptoms

Disposition of Public Comments

Commenter:

Drew B. Schembre, MD, Gastroenterologist, Swedish, Seattle, WA, *HERC-appointed Expert*

Comments Grouped by Topic

#	Comment	Disposition
1	Here are some recent guidelines from three separate GI societies regarding diagnostic approaches to GERD and dyspepsia. There are well reasoned decision trees that are worth reviewing. There are additional cost analyses that are interesting. I believe these support my contention that recognition of alarm symptoms, age > 50 and persistent GERD symptoms after a 2 week trial of treatment warrant upper endoscopy.	Thank you for this additional information.
2	American Society for Gastrointestinal Endoscopy. (2007). Role of endoscopy in the management of GERD. <i>Gastrointestinal Endoscopy</i> , 66(2), 219-224. <i>Rated poor quality by CEBP¹</i>	<p>Recommendations pertinent to this guidance include the following:</p> <ul style="list-style-type: none"> “Endoscopy is recommended for patients who have symptoms suggesting complicated GERD or alarm symptoms (2A).” <p>2A recommendation is described as intermediate strength; unclear benefit; based on RCTs</p> <p>Alarm symptoms are listed as GERD symptoms persistent or progressive despite medical therapy (length of therapy not specified), dysphagia/odynophagia, involuntary weight loss (>5%), GI bleeding/anemia, presence of mass/stricture/ulcer, persistent vomiting, suspected extra-esophageal manifestations of GERD (latter indication contradicted by Katz and Kahrilas).</p> <ul style="list-style-type: none"> “Endoscopy should be considered in patients at risk for Barrett’s esophagus (BE) (level of evidence = 2C).” <p>2C recommendation is described as very weak recommendation, alternative approaches likely to be better under some circumstances; unclear benefits; based on observational studies.</p>

¹ The Center for Evidence-based Policy (CEbP) assesses the methodological quality of guidelines using an instrument adapted from the Appraisal of Guidelines Research and Evaluation (AGREE) Collaboration (<http://www.agreetrust.org/resource-centre/practice-guidelines/>). Guidelines are assigned a rating of good, fair, poor, based on its adherence to recommended methods and potential for biases.

HERC Coverage Guidance – Upper Endoscopy for Gastroesophageal Reflux Disease (GERD) and Dyspepsia Symptoms Disposition of Public Comments

#	Comment	Disposition
		<p>Risk factors for BE listed as prolonged (>5 years) GERD symptoms, white race, male sex, age > 50, + family history</p> <p><i>For HTAS discussion: what are appropriate alarm symptoms and/or indications for endoscopy in patients with GERD?</i></p>
3	<p>Katz, P.O., Gerson, L.B., & Vela, M.F. (2013). Guidelines for the Diagnosis and Management of Gastroesophageal Reflux Disease. <i>American Journal of Gastroenterology</i>, 108, 308-328. doi: 10.1038/ajg.2012.444</p> <p><i>Rated poor quality by CEbP</i></p>	<p>Recommendations pertinent to this guidance include the following:</p> <p>“Upper endoscopy is not required in the presence of typical GERD symptoms. Endoscopy is recommended in the presence of alarm symptoms and for screening of patients at high risk for complications. Repeat endoscopy is not indicated in patients without Barrett’s esophagus in the absence of new symptoms.”</p> <p>Alarm symptoms are specified as dysphagia, but not otherwise described. Those at high risk for complications are likewise not defined.</p> <p>“Upper endoscopy should be performed in refractory patients with typical or dyspeptic symptoms principally to exclude non-GERD etiologies.”</p> <p>Definition of refractory not provided, but author notes that poor compliance and inappropriate dosing are significant factors in lack of response to PPI and should be corrected first. Eight week course of PPIs is recommended as initial treatment.</p> <p>“Upper endoscopy is not recommended as a means to establish a diagnosis of GERD-related asthma, chronic cough, or laryngitis.”</p> <p><i>For HTAS discussion: is repeat endoscopy in patients with negative findings appropriate?</i></p>
4	<p>Talley, N.J., Vakil, N., & the Practice Parameters Committee of the American College of Gastroenterology. (2005). Guidelines for the Management of Dyspepsia. <i>American Journal of Gastroenterology</i>, 100, 2324-2337. doi: 10.1111/j.1572-0241.2005.00225.x</p> <p><i>Rated poor quality by CEbP</i></p>	<p>Recommendations pertinent to this guidance include the following:</p> <p>Patients with dyspepsia should undergo EGD if they are > 55 or have alarm symptoms, defined as:</p> <ul style="list-style-type: none"> • Bleeding or anemia • Early satiety • Unexplained weight loss > 10% of body weight • Progressive dysphagia or odynophagia • Persistent vomiting • Personal or family history of esophagogastric malignancy • History of peptic ulcer

HERC Coverage Guidance – Upper Endoscopy for Gastroesophageal Reflux Disease (GERD) and Dyspepsia Symptoms Disposition of Public Comments

#	Comment	Disposition
		<ul style="list-style-type: none"> Lymphadenopathy Abdominal mass <p>Repeat EGD is not recommended unless completely new symptoms or alarm features develop.</p> <p>Use of antisecretory therapy can mask a cancer at endoscopy, but does not appear to alter the outcome.</p> <p>Refractory GERD not defined, but recommendation is for initial 4-8 week course of PPI.</p> <p style="color: red;"><i>For HTAS discussion: should the age at which endoscopy is recommended be 55 rather than 50, and is repeat endoscopy in patients with negative findings appropriate?</i></p>
5	<p>Kahrilas, P.J., Shaheen, N.J., Vaezi, M.F., Hiltz, S.W., Black, E., Modlin, I.M., et al. (2008). American Gastroenterological Association medical position statement on the management of gastroesophageal reflux disease. <i>Gastroenterology</i>, 135(4), 1383-1391, 1391.e1-5.</p> <p><i>Rated good quality by CEbP</i></p>	<p>Recommendations pertinent to this guidance include the following:</p> <p><i>How Do Antisecretory Therapies Compare in Efficacy and Under What Circumstances Might One Be Preferable to Another? What Is an Acceptable Upper Limit of Empirical Therapy in Patients With Suspected Typical Esophageal GERD Syndromes Before Performing Esophagogastroduodenoscopy?</i></p> <p>PPIs are recommended for initial empiric treatment (Grade A). Authors state “Patients whose heartburn has not adequately responded to twice-daily PPI therapy should be considered treatment failures, making that a reasonable upper limit for empirical therapy.” However, length of initial trial of PPIs is not specified.</p> <p><i>What Is the Role and Priority of Diagnostic Tests (Endoscopy With or Without Biopsy, Esophageal Manometry, Ambulatory pH Monitoring, Impedance-pH Monitoring) in the Evaluation of Patients With Suspected Esophageal GERD Syndromes?</i></p> <p>Grade B: recommended with fair evidence that it improves important outcomes</p> <ol style="list-style-type: none"> Endoscopy with biopsy for patients with an esophageal GERD syndrome with troublesome dysphagia. Endoscopy to evaluate patients with a suspected esophageal GERD syndrome who have not responded to an empirical trial of twice-daily PPI therapy. <p>Grade Insuff: no recommendation, insufficient evidence to recommend for or against</p> <ol style="list-style-type: none"> Using alarm symptoms (other than troublesome dysphagia) as a screening tool to identify patients with GERD at risk for esophageal adenocarcinoma.

HERC Coverage Guidance – Upper Endoscopy for Gastroesophageal Reflux Disease (GERD) and Dyspepsia Symptoms Disposition of Public Comments

#	Comment	Disposition
		<p><i>Does GERD Progress in Severity, Such That Symptomatic Patients Without Esophagitis Develop Esophagitis and Barrett’s Metaplasia, or Are These Distinct Disease Manifestations That Do Not Exist Along a Continuum? If Patients Do Progress, at What Rate Does This Occur, and Does It Warrant Endoscopic Monitoring?</i></p> <p>Grade D: recommend against, fair evidence that it is ineffective or harms outweigh benefits I. Routine endoscopy in subjects with erosive or nonerosive reflux disease to assess for disease progression.</p> <p><i>What Is the Role of Endoscopy in Longterm Management of Patients With GERD, and Under What Circumstances Should Mucosal Biopsy Specimens Be Obtained When Endoscopy Is Performed?</i></p> <p>Grade B: recommended with fair evidence that it improves important outcomes I. Endoscopy with biopsy for patients with an esophageal GERD syndrome with troublesome dysphagia.</p> <p>Grade Insuff: no recommendation, insufficient evidence to recommend for or against I. Routine upper endoscopy in the setting of chronic GERD symptoms to diminish the risk of death from esophageal cancer. II. Endoscopic screening for Barrett’s esophagus and dysplasia in adults 50 years or older with >5–10 years of heartburn to reduce mortality from esophageal adenocarcinoma.</p> <p><i>For HTAS discussion: should endoscopy be recommended in patients over 50, or those with alarm symptoms other than dysphagia? Should repeat endoscopy be recommended in the absence of Barrett’s or dysplasia?</i></p>
6	<p>American Society for Gastrointestinal Endoscopy. (2012). The role of endoscopy in Barrett’s esophagus and other premalignant conditions of the esophagus. <i>Gastrointestinal Endoscopy</i>, 76(6), 1087-1094.</p> <p><i>Rated poor quality by CEbP</i></p>	<p>Recommendations pertinent to this guidance include the following:</p> <ol style="list-style-type: none"> Endoscopic screening for BE can be considered in select patients with multiple risk factors for Barrett’s esophagus (BE) and esophageal adenocarcinoma (EAC), but patients should be informed that there is insufficient evidence to affirm that this practice prevents cancer or prolongs life. Risk factors are defined as male sex, white race, age > 50, + family history, increased duration of reflux symptoms, smoking and obesity. We recommend no further endoscopic screening for BE after a screening examination with negative findings. <p><i>For HTAS discussion: is repeat endoscopy in patient with negative findings appropriate?</i></p>

HEALTH EVIDENCE REVIEW COMMISSION (HERC)

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

DRAFT for HTAS Meeting Materials 9/23/2013

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

- Represents high costs, significant economic impact
- Topic is of high public interest

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

EVIDENCE SOURCE

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/uspstf.htm](http://www.uspreventiveservicestaskforce.org/uspstf/uspstf/uspstf.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,

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 osteoporosis¹,
 - 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 FRAX² (without a bone mineral density [BMD] value, if a dual-energy X-ray absorptiometry [DXA] scan has not previously been undertaken) or QFracture³, 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 threshold⁴ 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 threshold⁵ 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)

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

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.

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.

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

Appendix C. HERC Guidance Development Framework

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

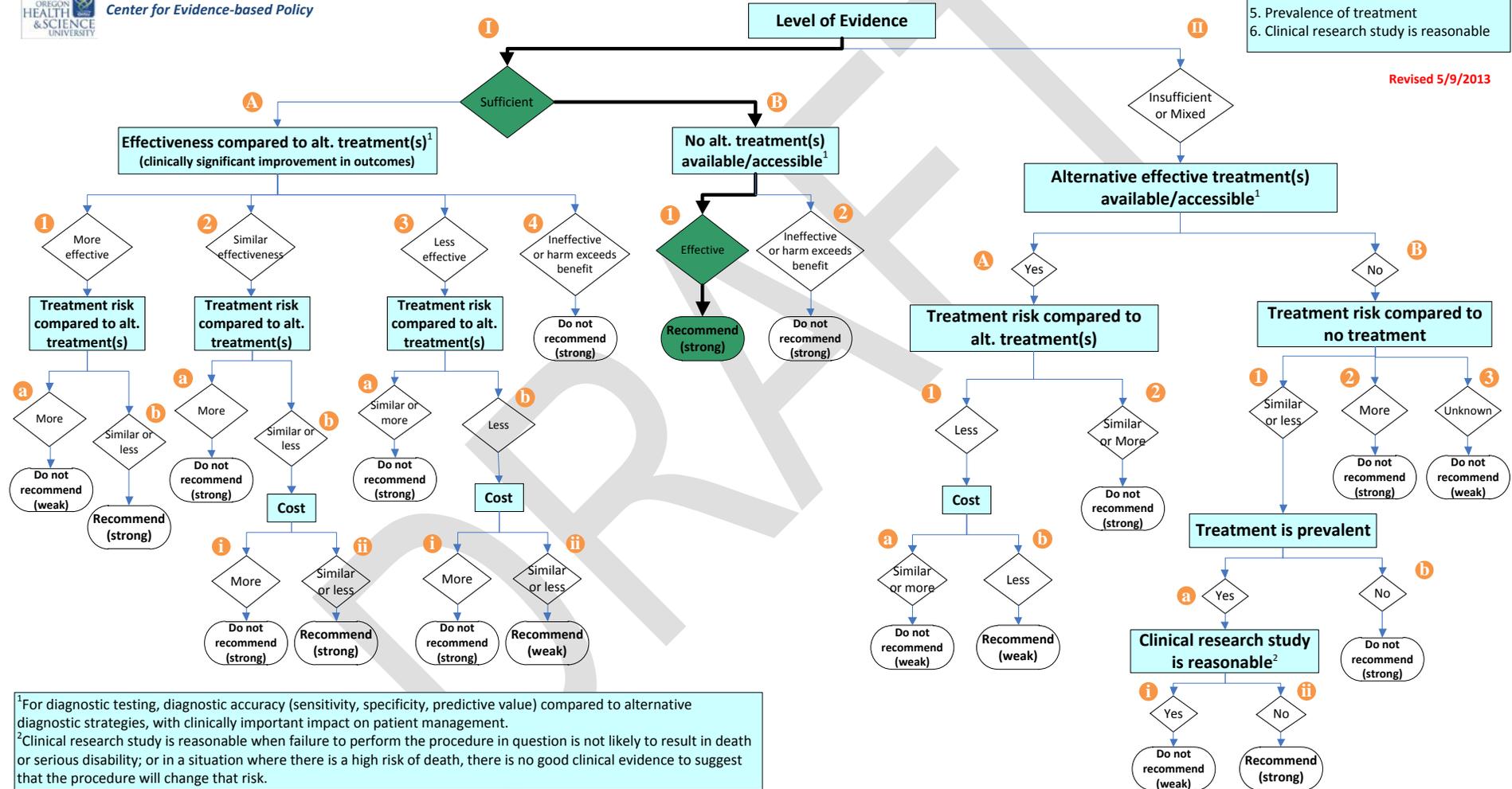


HERC Guidance Development Framework

Refer to HERC Guidance Development Framework Principles for additional considerations

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

Revised 5/9/2013



Screening for osteoporosis in men without additional risk factors

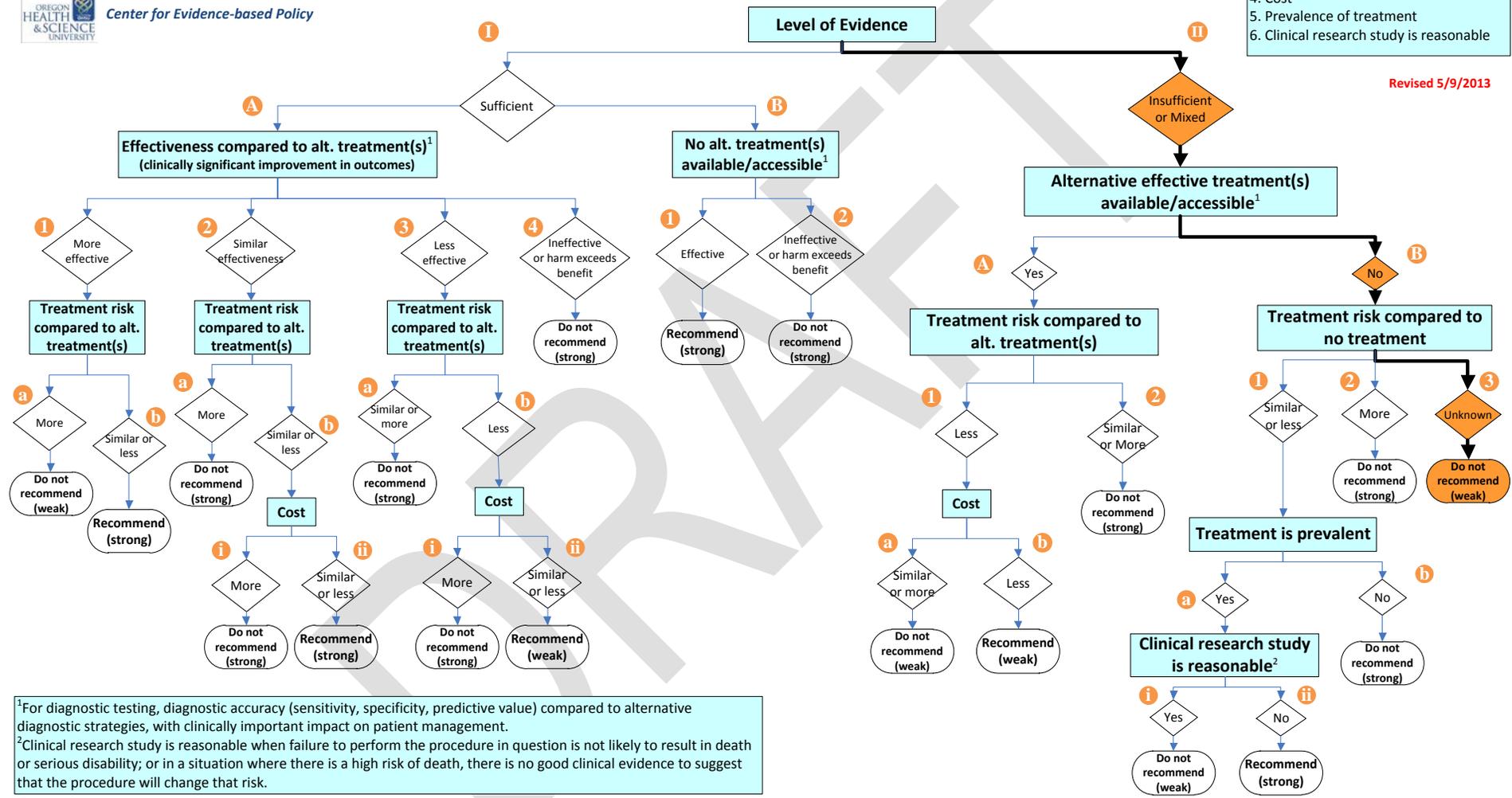


HERC Guidance Development Framework

Refer to HERC Guidance Development Framework Principles for additional considerations

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

Revised 5/9/2013



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



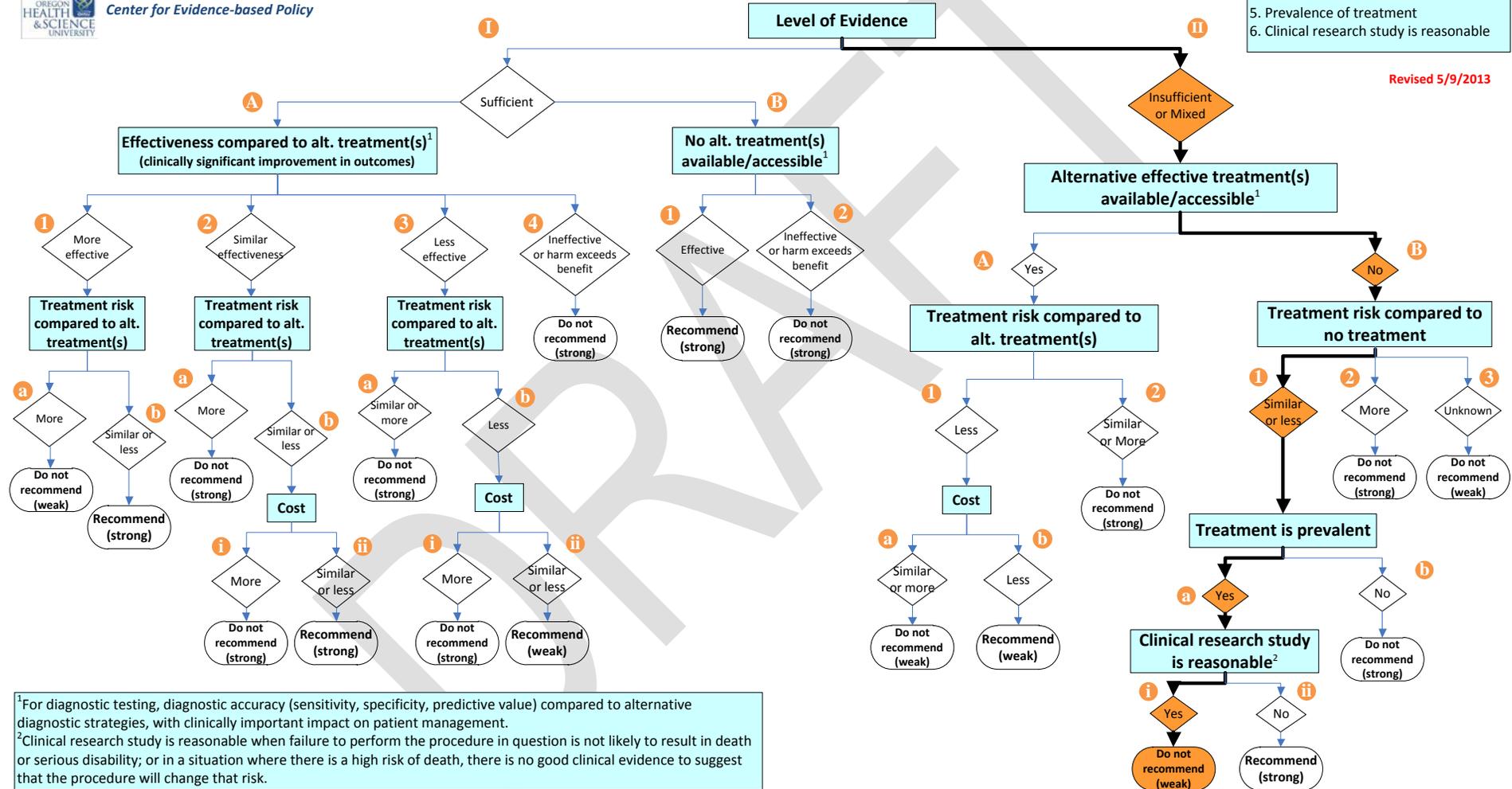
HERC Guidance Development Framework

Refer to HERC Guidance Development Framework Principles for additional considerations

- Decision Point Priorities**

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

Revised 5/9/2013



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

Disposition of Expert Comments

Expert	#	Comment	Disposition
E. Michael Lewiecki, MD, FACP, FACE Osteoporosis Director, New Mexico Clinical Research & Osteoporosis Center, Inc., Albuquerque, NM	1	<p>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.</p>	Thank you for this background information.
	2	<p>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:</p> <ul style="list-style-type: none"> • 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 \geq 5 mg prednisone or equivalent for \geq 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	<p>The NOF guidelines also describe the use of DXA to monitor osteoporosis therapy, as follows:</p> <ul style="list-style-type: none"> • 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

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

Disposition of Expert Comments

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.	measure a change in BMD; 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 rescreeing 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

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

Disposition of Expert Comments

Expert	#	Comment	Disposition
			should undergo screening if they have at least one major or two minor risk factors for osteoporosis.” Ref #19 is also a CEA 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	HTAS is aware that the USPSTF does not address the use

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

Disposition of Expert Comments

Expert	#	Comment	Disposition
		of DXA other than screening (e.g., monitoring) or DXA in men. It should be noted that 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.	of DXA in monitoring, and therefore includes the Gourlay study in the guidance document to address this void. The USPSTF <u>does</u> 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	Ref #27 is a prospective case series of 671 postmenopausal 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

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

Disposition of Expert Comments

Expert	#	Comment	Disposition
		Gourlay et al.	was also increased with age, prior fracture and high bone turnover 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	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.

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

Disposition of Expert Comments

Expert	#	Comment	Disposition
		use in their publication but has no established definition.	
	15	<p>Reference List</p> <ol style="list-style-type: none"> 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. 	

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

Disposition of Expert Comments

Expert	#	Comment	Disposition
	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.	
	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.	

**HERC Coverage Guidance – Osteoporosis Screening And Monitoring By
Dual-Energy X-Ray Absorptiometry (DXA)
Disposition of Expert Comments**

Expert	#	Comment	Disposition
		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 9/23/2013

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:

CHOOSE Option 1, 2, or 3:

Option 1 (aligns with Medicare, previously approved by HTAS)

- 12 week 'trial' period to determine benefit. This period is covered if apnea-hypopnea index (AHI) or respiratory disturbance index (RDI) is greater than or equal to **15** events per hour, or if between **5 and 14 events** with additional symptoms including excessive daytime sleepiness (Epworth Sleepiness Scale score > 10), or documented hypertension, ischemic heart disease, or history of stroke;

Option 2 (cutoff of 15 aligns with improvement in sleepiness and QOL shown in some studies)

- 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 excessive daytime sleepiness (Epworth Sleepiness Scale score > 10), or documented hypertension, ischemic heart disease, or history of stroke;

Option 3 (aligns with mortality benefit in non-treatment studies)

- 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.

The following bullets would apply to any of the three options above

- 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 is not recommended for coverage ~~only covered after a diagnosis of sleep apnea has been made, and there is documented failure or intolerance of both CPAP and an oral appliance (or other non-invasive treatment), and patients have been informed of the benefits and risks of surgery.~~

RATIONALE FOR GUIDANCE DEVELOPMENT

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

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

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

EVIDENCE SOURCE

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 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.

~~The AHI has variable value as a predictor of clinical outcomes:~~

- ~~● The strength of evidence is high (based on four trials) that high baseline (AHI>30 events/hr or range) AHI is a strong and independent predictor of all-cause mortality over several years of follow-up (2-14 years).~~
- ~~● The association between baseline AHI and the other long-term clinical outcomes is less robust, having been analyzed by only one or two studies:~~
 - ~~○ Cardiovascular (CV) disease (studies reported mixed results regarding CV death, but AHI >30 was an independent predictor of nonfatal CV disease.~~
 - ~~○ Stroke (one study suggested that the association between AHI and stroke may be confounded by obesity).~~
 - ~~○ Hypertension (studies had uncertain conclusions regarding the possible association between AHI and incident hypertension)~~
 - ~~○ Non-insulin-dependent diabetes and other metabolic abnormalities (studies reported mixed results that suggested an association between AHI and incident type 2 diabetes which, in one study, was confounded by obesity)~~
 - ~~○ Decreased quality of life (a single study found no significant association between AHI and future quality of life [SF-36 after 5 years]).~~
 - ~~○ No current established threshold level for AHI exists that indicates the need for treatment.~~

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 Scale¹ (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 (AHI), improving symptoms as measured by the ESS, reducing arousal index, and raising the minimum oxygen saturation). Weak evidence demonstrated no consistent

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.

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 \geq 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

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.](#)

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
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

CODES	DESCRIPTION
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	
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 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 inconvenience.	Moderate	Moderately costly	Moderate variability	

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
CPAP for patients with AHI \geq 30	Significant benefit on mortality/ comorbid diseases, moderate benefit on sleepiness/QOL. No serious harms, but significant patient inconvenience.	Moderate	Moderately costly	Small variability	

*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.

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

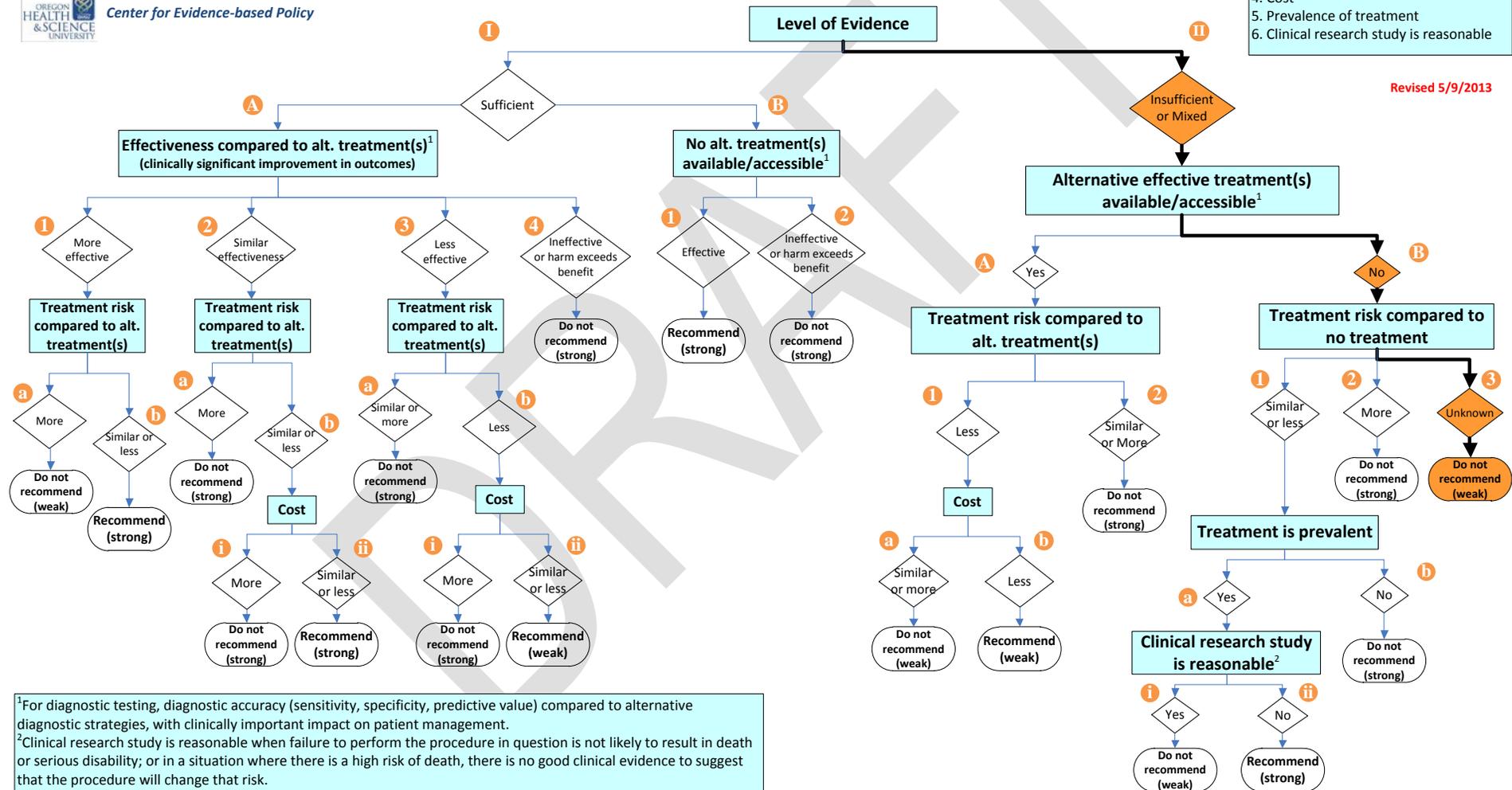


HERC Guidance Development Framework

Refer to HERC Guidance Development Framework Principles for additional considerations

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

Revised 5/9/2013



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

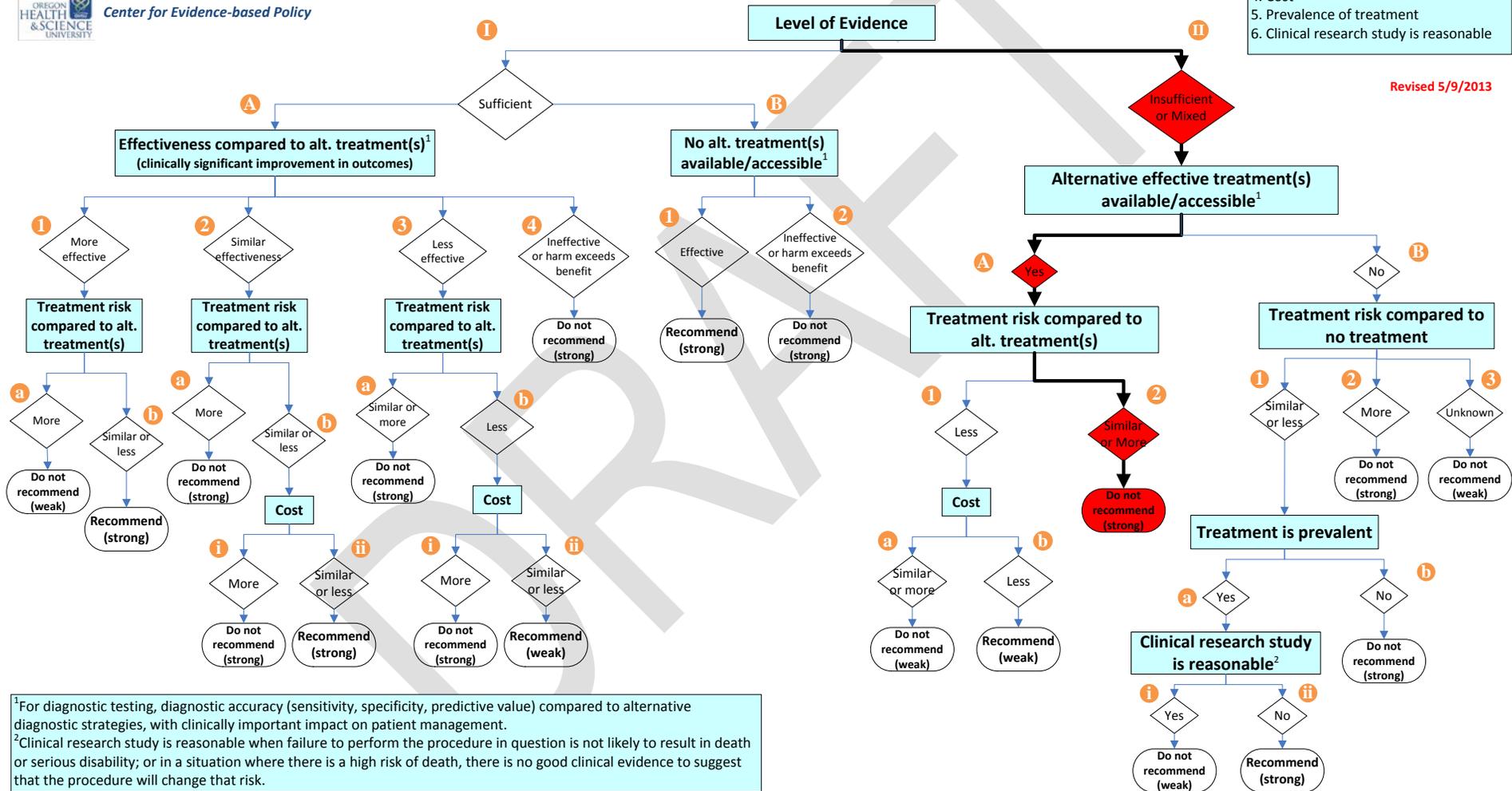


HERC Guidance Development Framework

Refer to HERC Guidance Development Framework Principles for additional considerations

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

Revised 5/9/2013



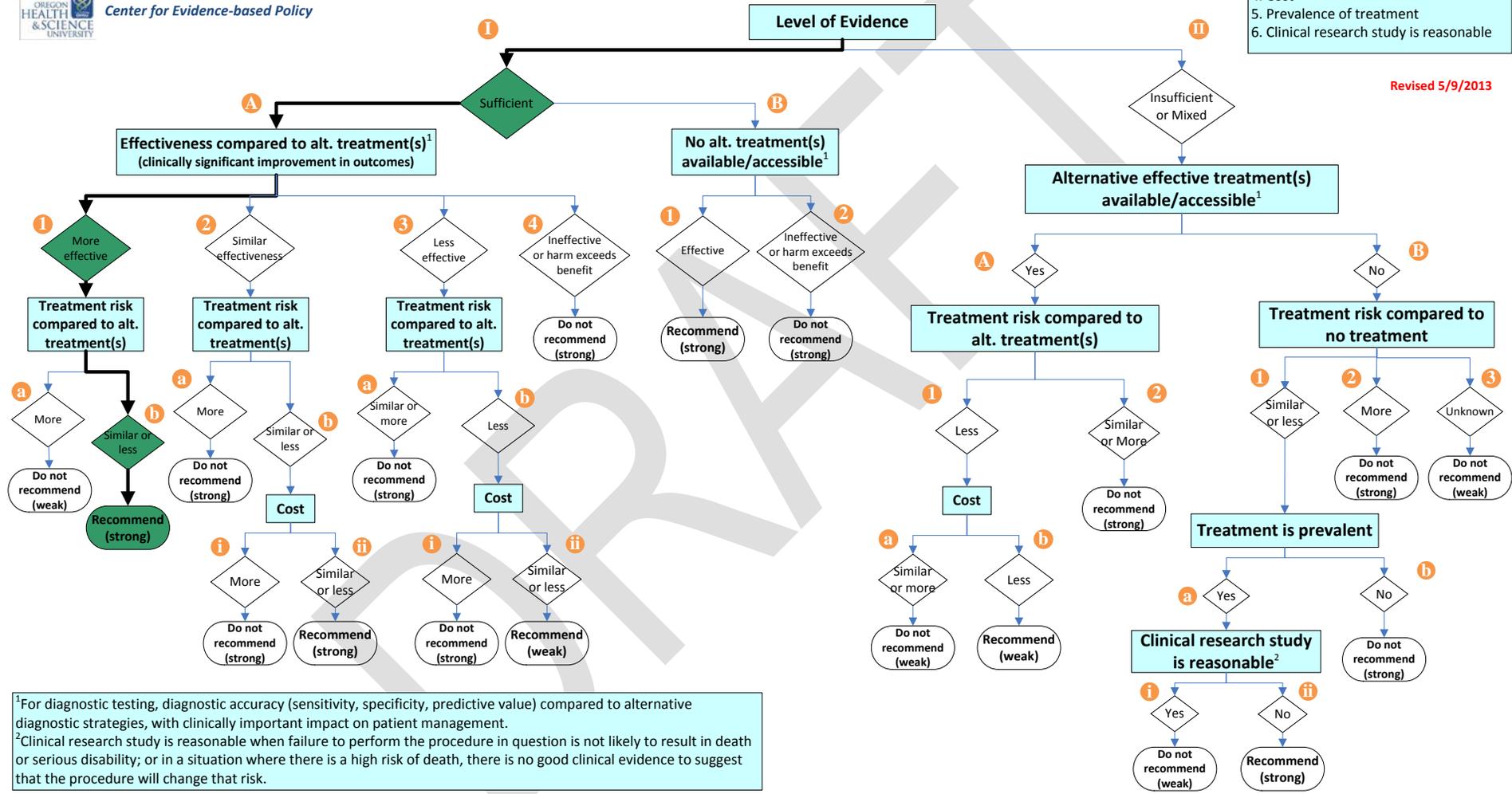
CPAP for Patients with AHI 15-29; CPAP for Patients with AHI ≥ 30



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

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

Revised 5/9/2013



¹For diagnostic testing, diagnostic accuracy (sensitivity, specificity, predictive value) compared to alternative diagnostic strategies, with clinically important impact on patient management.
²Clinical research study is reasonable when failure to perform the procedure in question is not likely to result in death or serious disability; or in a situation where there is a high risk of death, there is no good clinical evidence to suggest that the procedure will change that risk.

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

General Comments

Stakeholder	#	Comment	Disposition
<i>Medical Director, Health Plan</i> 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.

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

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.