Health Evidence Review Commission's Evidence-based Guideline Subcommittee

June 5, 2014

2:00 PM

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

Section 1.0

Call to Order

AGENDA

EVIDENCE-BASED GUIDELINES SUBCOMMITTEE (EbGS)
June 5, 2014
2:00pm - 5:00pm
Meridian Park Hospital
Community Health Education Center Room 117B&C

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

#	Time	Item	Presenter
1	2:00 PM	Call to Order	Wiley Chan
2	2:05 PM	Review of April 24, 2014 minutes	Wiley Chan
3	2:10 PM	Staff report	Cat Livingston Darren Coffman
4	2:15 PM	New Draft Coverage Guidance: · Ablation for Atrial Fibrillation · Public comment (5 minutes)	Cat Livingston
5	3:30 PM	New Draft Coverage Guidance: · Nuclear Cardiac Imaging · Public comment (5 minutes)	Cat Livingston
6	4:45 PM	Public Comment	
7	4:55 PM	Confirmation of next meeting (September 4, 2014)	Wiley Chan
9	5:00 PM	Adjournment	Wiley Chan

MINUTES

Evidence-based Guidelines Subcommittee
Oregon Dental Association
Conference Center
8699 SW Sun Place
Wilsonville, OR 97070
April 24, 2014
2:00-5:00pm

Members Present: Wiley Chan, MD, Chair; Vice-Chair; Vern Saboe, DC; Beth Westbrook, PsyD; John Sattenspiel, MD, MPH; Bob Joondeph, JD; Eric Stecker, MD, MPH (participated by phone from approximately 3:15-4:30).

Members Absent: Steve Marks, MD; Leda Garside, RN, MBA; Som Saha, MD, MPH;

Staff Present: Darren Coffman; Cat Livingston, MD, MPH; Jason Gingerich.

Also Attending: Alison Little, MD (CEBP); Shannon Vandergriff (CEBP); Paul Terdal (Autism Speaks), Barbara Morrow (Astellas).

---

1. CALL TO ORDER

The meeting lacked a quorum until 2:30 p.m. Wiley Chan called the meeting of the Evidence-based Guidelines Subcommittee (EbGS) to order at 2:30 pm.

---

2. MINUTES REVIEW

No changes were made to the March 20, 204 minutes.
Minutes approved 5-0 (Stecker not present).

---

3. EVALUATION OF EVIDENCE: APPLIED BEHAVIOR ANALYSIS (ABA) FOR AUTISM SPECTRUM DISORDERS

A) Update on ABA process and implementation

Coffman provided a review of the process to date and said that in order to meet the legislative timeline, the EbGS needs to approve the document today. Chan invited public testimony.

B) Public comment

Paul Terdal gave public testimony and also introduced Vera, who spoke about her experience with her son, who will soon be thirteen and who has an autism spectrum disorder and is covered under OHP. She said that her son has symptoms including insomnia, toe-walking, hair pulling and he is almost entirely nonverbal. She believes he

DRAFT

has feeling and thoughts but cannot express them. She is concerned about the proposed limits on therapy for children over the age of 12, because she is hoping that expanded coverage of ABA can make a difference for her son.

Terdal said that he appreciates the work of the subcommittee but disagrees with the subcommitee's choices about the sufficiency of evidence and reasonableness of a trial in the population of children over the age of 12. He referenced several studies referenced in the subcommittee's evidence sources which did include some children over 12. He said there is no evidence for decline of effectiveness by age. Changing either of these two factors would change the recommendation according to the HERC Guidance Development framework to a strong recommendation for coverage. He also referenced legal decisions about the reasonableness of a trial. He said that Vera's child would not be here to day if he had received therapy, either because ABA had been effective or because it hadn't and he had sought other treatments.

C) Discussion and approval of modified Evidence Review to send to VbBS/HERC

Livingston reviewed the changes since the last meeting, as described in the meeting materials on page 12 of the PDF document. During the discussion, the group asked that the remaining references to the age range of 2 to 12 be changed to 1 to 12 per the discussion at the subcommittee's previous meeting. In addition, the group decided to use the proposed new pathway in the coverage guidance development framework for ABA treatments for children aged 1 to 12, including EIBI and other less intensive interventions, though the recommendation is still a strong recommendation for coverage. In addition, the group decided to change the language in the GRADE-informed framework from "Comprehensive ABA" to Early Intensive Behavioral Interventions.

For adolescents and young adults, the group made no changes to the guidance development framework, though Joondeph expressed concern that the cutoff was arbitrary. Sattenspiel said that without evidence of global benefit for the broader therapy, the limited coverage shown in this draft makes sense for older children. Westbrook said she also struggles with how it may feel unfair, but said we need to be consistent with how we've treated the evidence in other reviews. Chan noted that in medicine, similar decisions are frequently made based on inclusion criteria for studies. For instance, the use of statins is not well-studied in patients over the age of 75. Joondeph said that he finds the single-subject research design studies submitted by members of the public compelling, even if they do not meet the Commission's normal criteria for sufficient evidence. Westbrook noted, however, that many of these studies used different populations, including people without autism spectrum disorders, and that SSRDs are subject to bias.

After more discussion the group decided to ask Livingston to add language reflecting the subcommittee's differing opinions about the evidence for older individuals.

The group then moved on to discuss the summary conclusions. Though there is no good evidence on which to base decisions about limits on the number of hours, the group decided to use a maximum intensity of 25 hours per week for a maximum of 3 years, using the rationales from the meeting materials.

Larsson expressed concern that these numbers were based on averages and may not accommodate the needs of some patients. His recommendation for coverage for EIBI,

DRAFT

submitted with the packet materials, recommended an average of 20 hours of behavior technician time and an average of 7 hours of behavior analyst services for a maximum of 3 years. Joondeph said that setting a limit based on the average from studies didn't make sense to him. Livingston explained that the evidence included studies with various intensities, above and below 25 hours, and that we lack evidence that the more intensive interventions are more effective. The 25 hour limit would be the same as commercial payers. She said she sees no evidence against a 25 hour limit and does see some reasons to support it. Westbrook asked for clarifications about the exceptions and appeals processes, and Sattenspiel explained that CCOs are required to allow for exceptional circumstances, but are not reimbursed for providing those services. If CCOs deny coverage, patients have the right to appeal the decision.

After reviewing an emailed comment provided by Dr. Fombonne, an appointed expert for this review, the group decided to change the references to "comprehensive ABA" to reference EIBI and clarify that the UCLA Lovaas and ESDM models referenced are examples. The listed standardized assessments are also examples.

The group also discussed the language for less intensive ABA-based interventions, which has been added to the draft. The subcommittee changed the recommendation for duration to allow for an initial six months of coverage, with ongoing coverage based on demonstrated progress towards meaningful predefined objectives or the emergence of new problem behaviors.

For individuals 13 and older the subcommittee clarified the rationale regarding the need for meaningful progress towards predefined goals and changed the phrase "medical necessity" to "medical appropriateness." In addition, the group clarified that coverage for training of parents and other caregivers is appropriate.

The group also made changes to the duration and frequency after broad agreement that reauthorizations more often than every six months don't make sense, as assessment and post-treatment evaluation would take up some of the time.

After further discussion, the group voted to approve the draft evaluation of evidence as modified during the meeting and with additional language by Livingston to reflect disagreement among subcommittee members about strength of evidence in children over the age of 12. **Motion approved 5-0** (Stecker absent)

---

4. REVIEW OF NEW DRAFT COVERAGE GUIDANCES ON NUCLEAR CARDIAC IMAGING

Discussion deferred to the next meeting as there was no time to discuss the topic.

---

5. PUBLIC COMMENT

Chan invited additional public comment, but no one chose to testify.

DRAFT

6. ADJOURNMENT

The meeting was adjourned at 5:00 pm. The next meeting is scheduled for June 5, 2014 from 2:00-5:00pm in Room 117B/C of the Meridian Park Hospital Community Health Education Center in Tualatin.

DRAFT

Section 2.0

Coverage Guidances

HEALTH EVIDENCE REVIEW COMMISSION (HERC)

COVERAGE GUIDANCE: ABLATION FOR ATRIAL FIBRILLATION

DRAFT for June 5, 2014 EbGS meeting materials

HERC COVERAGE GUIDANCE

AV node ablation is recommended for coverage only in symptomatic persons when pharmacological therapy for rate control is ineffective or not tolerated (weak recommendation)

Pulmonary vein isolation is recommended for coverage with (weak recommendation) or without other cardiac surgery (strong recommendation) *for those who are persistently symptomatic despite rate control medications +/- antiarrhythmics*

The Maze procedure is recommended for coverage at the time of other cardiac surgery if the benefits of maintenance of sinus rhythm are thought to outweigh the potential risk of increased mortality *(weak recommendation)*

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

RATIONALE FOR GUIDANCE DEVELOPMENT

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

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

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

EVIDENCE SOURCES

Al-Khatib, S.M., Allen Lapointe, N., Chatterjee, R., Crowley, M.J., Dupre, M.E., Kong, D.F., et al. (2013). *Treatment of atrial fibrillation. Comparative Effectiveness Review 119*. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2007-10066-I.) AHRQ Publication No.13-EHC095-EF.

DRAFT

Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from www.effectivehealthcare.ahrq.gov/reports/final.cfm

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

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

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

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

SUMMARY OF EVIDENCE

Clinical Background

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

It is estimated that more than 2.3 million Americans have AF. The prevalence of AF increases with age and approaches 8 percent in patients older than 80 years of age. AF is the most common sustained arrhythmia seen in clinical practice. The impact of AF is compounded by its known association with significant mortality, morbidity, and health care costs. Not only is the risk of death in patients with AF twice that of patients without AF, but AF can result in myocardial ischemia or even infarction, heart failure

exacerbation, and tachycardia-induced cardiomyopathy if the ventricular rate is not well controlled. The most dreaded complication of AF is thromboembolism, especially stroke. The risk of stroke in patients with AF is up to 8 percent per year, depending on the presence of stroke risk factors. Importantly, when ischemic stroke occurs in patients with AF, it is either fatal or of moderate to high severity in the majority of patients. The management of AF and its complications is responsible for almost $16 billion in costs to the U.S. health care system each year.

Treatment Strategies

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

Rate Control

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

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

Rhythm Control

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

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

Several other procedures for the treatment of AF have been investigated. One such procedure is the surgical Maze procedure, which appears to confer some benefit to selected patients with AF. Implantation of a cardiac resynchronization therapy (CRT) device is another procedure that may decrease the burden of AF in patients who are eligible for this device based on a left ventricular ejection fraction ≤35 percent, a wide QRS complex, and heart failure symptoms despite optimal medical therapy. Secondary analyses of major clinical trials have provided conflicting findings on the effect of CRT on AF burden.

Although several studies of rate- and rhythm-control strategies exist, to date no study has shown that maintaining patients with AF in sinus rhythm provides a long-term survival benefit. It is also unknown whether the risks and benefits of different therapies vary by AF type.

Evidence Review

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

Al-Khatib 2013 reports on six RCTs (two good, three fair, and one poor quality) involving a total of 537 patients that compared the effectiveness of a procedural intervention versus a primarily pharmacological intervention for rate control of AF or comparing two primarily procedural interventions. Four studies recruited patients with only one specific type of AF, either permanent (three studies) or persistent (one study); one study

recruited patients with “resistant chronic” AF; and one study recruited patients with permanent or paroxysmal AF. These studies, however, evaluated and compared different types of treatments, preventing conclusions about whether effectiveness varied by type of AF. All studies included at least one treatment arm with radiofrequency ablation of either the AVN or His bundle, most often in conjunction with pacemaker placement. The comparison arms included a pharmacological intervention whose main purpose was to control ventricular heart rate rather than converting the underlying rhythm of AF; this was combined with a procedure in some studies.

Based on three studies reported in Al-Khatib 2013 (one good, two poor quality) involving 175 patients, patients undergoing a procedural intervention had a significantly lower heart rate at 12 months than those receiving a primarily pharmacological intervention. This was measured differently in all three studies. In one, the mean heart rate in the intervention group was 71 ±6 bpm compared to 83 ± 8 bpm in the medication group ($p<0.01$). In this study, maximum heart rate did not differ between groups. In the second study, those in the ablation group had higher minimum (70±9 vs. 39±9 bpm; $p<0.05$) and mean (76±7 vs. 71±11 bpm; $p<0.05$) heart rates than the medication group, but lower maximum heart rates (117 ± 16 bpm vs. 152 ± 37 bpm; $p<0.05$). The third study the percent of each group who had either a normal or uncontrolled ventricular rate; in the ablation group, 100% had a normal ventricular rate (50-90 bpm) compared to 58% in the medication group. Similarly, none of the ablation group had an uncontrolled heart rate (>90 bpm at rest or > 130 bpm on exertion), while 42% of the medication group did. There was no difference by treatment arm in all-cause mortality (two studies [one good, one fair quality], 201 patients); cardiovascular mortality (one study [good quality], 102 patients); or exercise capacity (two studies [one good, one fair quality], 135 patients) (all low strength of evidence). There was insufficient strength of evidence to support findings for other outcomes, including quality of life.

<u>Rhythm-Control Procedures and Drugs for Maintenance of Sinus Rhythm</u>

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

Transcatheter PVI versus antiarrhythmic drugs

Al-Khatib 2013 concluded, based on eight RCTs (five good, three fair quality) involving 921 patients, that transcatheter PVI is superior to antiarrhythmic drugs for **maintenance of sinus rhythm** over 12 months of follow up in patients with AF (one RCT reported 48 months of follow up). All trials had statistically significant results, as did meta-analysis of all eight trials (OR 6.51, 95% CI 3.22 to 13.16). This evidence is strongest in younger

patients with little to no structural heart disease and with mild or no enlargement of the left atrium. Only one trial was limited to patients receiving ablation as first line therapy (Wazni 2005), while five specifically required failure of at least one AAD to be included in the study. The Wazni trial included 70 patients who experienced monthly episodes of symptomatic AF for at least three months, and found that at one year follow up, 63% of those treated with AADs had at least one recurrence of AF, compared to 13% of those who received PVI. Another trial included only in Hashimoto 2013 included only patients with persistent AF (MacDonald 2011), and reported that at final follow up (6 months), 50% of patients in the PVI group were in sinus rhythm while none of the control group were (no statistical testing done). This latter trial was limited to patients with advanced heart failure. (Note: This outcome is reported as freedom from recurrence in Hashimoto 2013, but results are similar.)

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

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

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

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

0.001) or at a mean follow up of 69 months (Sonne 2009: 2.1% in the PVI group vs. 16.5% in the AAD group, p = 0.001).

Eight studies evaluated **quality of life (QOL) or functional status**, three RCTs reported in all three source reports, two additional RCTs reported in both Hashimoto 2013 and Al-Khatib 2013, two additional RCTs in Hashimoto 2013 only and one cohort study reported in Al-Khatib 2013 only. In general, there was greater improvement from baseline in these scores in patients randomized to the PVI arm, compared to the AAD arm, and in most of these studies, results were statistically significant for at least some measures.

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

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

TE = thromboembolic; PV = pulmonary vein

Cryoablation PVI vs. AAD

One RCT reported in Hashimoto 2013 found that patients randomized to receive cryoablation had significantly greater freedom from recurrence compared with those

patients randomized to receive AADs alone (69.9% versus 7%, respectively; P < .001). There was one death (0.6% in the cryoablation PVI group and none in the AAD group at 12 months, which was not statistically significant.

Surgical Maze versus standard of care (mitral valve surgery)

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

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

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

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

Other comparisons

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

- Transcatheter PVI using different types of ablation catheters
- Transcatheter PVI with Cavotricuspid isthmus (CTI) ablation vs. transcatheter PVI without CTI ablation
- Transcatheter PVI vs transcatheter PVI with ablation sites other than CTI and CFAE and transcatheter PVI involving all four PVs vs transcatheter PVI involving arrhythmogenic PVs only
- Transcatheter PVI Alone vs transcatheter PVI plus postablation antiarrhythmic drugs

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

Rate- Versus Rhythm-Control Therapies

There was insufficient strength of evidence about outcomes comparing a rhythm-control strategy that involved PVI with a rate-control strategy that involved AVN ablation and implantation of a pacemaker (one good-quality study) or rate-controlling medications (one poor-quality study).

[Evidence Source]

Evidence Summary

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

The surgical Maze procedure, when done at the time of other cardiac surgery, results in a higher likelihood of maintaining sinus rhythm than not performing the Maze (moderate SOE). Similarly, PVI done at the time of other cardiac surgery results in a higher likelihood of maintaining sinus rhythm than not performing PVI (high SOE), and no apparent difference in all-cause mortality or stroke (low SOE). The evidence is insufficient to assess the comparative benefits of a rhythm-control strategy that involves PVI with a rate-control strategy that involves AVN ablation or rate-controlling medications.

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

DRAFT

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variability in values and preferences	Coverage recommendation	Rationale
	hospitalizations, better QOL, possibly lower mortality	8 fair to good quality studies, depending on the outcome			recommendation)	drugs alone
Maze procedure	Better maintenance of SR; possible (nonsignificant) increase in mortality	Moderate based on 1 good and six fair quality studies	Moderate (concurrent with other cardiac surgery)	Moderate	The Maze procedure is recommended for coverage at the time of other cardiac surgery if the benefits of maintenance of sinus rhythm are thought to outweigh the potential risk of increased mortality *(weak recommendation)*	Maze may help maintain sinus rhythm but concerning nonsignificant increased risk of mortality
PVI done with other cardiac surgery	Better maintenance of SR	High based on 5 good and 3 fair quality studies	Moderate (concurrent with other cardiac surgery)	Low	PVI is recommended for coverage *(weak recommendation)*	For EbGS discussion
Rate vs. rhythm control therapies that include PVI or AV node ablation	Unknown	Insufficient (1 good and 1 poor quality study)	High	Moderate	For EbGS discussion	For EbGS discussion

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

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

Note: GRADE framework elements are described in Appendix A

DRAFT

POLICY LANDSCAPE

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

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

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

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

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

They cite the 2011 publication of the ACCF/AGA guidelines on the management of patient with AF as supporting evidence. These guidelines were recently updated (2014), and are rated fair quality using the MED standard criteria, primarily because study selection criteria was not specified and the quality of included studies was not assessed. These guidelines state the following with regard to AV node ablation for rate control in AF:

> Class IIa
> 3. AV nodal ablation with permanent ventricular pacing is reasonable to control the heart rate when pharmacological therapy is inadequate and rhythm control is not achievable. *(Level of Evidence: B)*
>
> Class III: Harm
> 1. AV nodal ablation with permanent ventricular pacing should not be performed to improve rate control without prior attempts to achieve rate control with medications. *(Level of Evidence: C)*

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

Class I

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

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

Class IIa

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

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

Class IIb

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

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

Class III: Harm

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

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

<u>COMMITTEE DELIBERATIONS – VALUE-BASED BENEFITS SUBCOMMITTEE</u>

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

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

Appendix A. GRADE Element Descriptions

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

Strong recommendation

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

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

Weak recommendation

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

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

Quality or strength of evidence rating across studies for the treatment/outcome[1]

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

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

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

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

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

DRAFT

Appendix B. Applicable Codes

CODES	DESCRIPTION
ICD-9 Diagnosis Codes	
427.31	Atrial fibrillation
ICD-10 Diagnosis Codes	
I48.0	Paroxysmal atrial fibrillation
I48.1	Persistent atrial fibrillation
I48.2	Chronic atrial fibrillation
I48.91	Unspecified atrial fibrillation
ICD-9 Volume 3 (Procedure Codes)	
None	
CPT Codes	
33250	Operative ablation of supraventricular arrhythmogenic focus or pathway (eg, Wolff-Parkinson-White, atrioventricular node re-entry), tract(s) and/or focus (foci); without cardiopulmonary bypass (For intraoperative pacing and mapping by a separate provider, use 93631) Codes 33254-33256 are only to be reported when there is no concurrently performed procedure that requires median sternotomy or cardiopulmonary bypass.
33251	…with cardiopulmonary bypass
33254	Operative tissue ablation and reconstruction of atria, limited (eg, modified maze procedure)
33255	Operative tissue ablation and reconstruction of atria, extensive (eg, maze procedure); without cardiopulmonary bypass
33256	…with cardiopulmonary bypass
33257	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), limited (eg, modified maze procedure) (List separately in addition to code for primary procedure)
33258	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), extensive (eg, maze procedure), without cardiopulmonary bypass (List separately in addition to code for primary procedure)
33259	Operative tissue ablation and reconstruction of atria, performed at the time of other cardiac procedure(s), extensive (eg, maze procedure), with cardiopulmonary bypass (List separately in addition to code for primary procedure)
33261	Operative ablation of ventricular arrhythmogenic focus with cardiopulmonary bypass
33265	Endoscopy, surgical; operative tissue ablation and reconstruction of atria, limited (eg, modified maze procedure), without cardiopulmonary bypass
33266	…operative tissue ablation and reconstruction of atria, extensive (eg, modified maze procedure), without cardiopulmonary bypass
93613	Intracardiac electrophysicologic 3-dimensional mapping (List separately in addition to code for primary procedure)
93650	Intracardiac catheter ablation of atrioventricular node function, atrioventricular conduction for creation of complete heart block, with or without temporary pacemaker placement
93653	Comprehensive electrophysiologic evaluation including insertion and repositioning of multiple electrode catheters with induction or attempted induction of an arrhythmia with right atrial pacing and recording, right ventricular pacing and recording, His recording with intracardiac catheter ablation of arrhythmogenic focus; with treatment of supraventricular tachycardia by ablation of fast or slow atrioventricular pathway, accessory atrioventricular connection, cavo-tricuspid isthmus or other single atrial focus or source of atrial re-entry (Do not report 93653 in conjunction with 93600-93603, 93610, 93612, 93618-93620, 93642, 93654)

DRAFT

CODES	DESCRIPTION
93655	Intracardiac catheter ablation of a discrete mechanism of arrhythmia which is distinct from the primary ablated mechanism, including repeat diagnostic maneuvers, to treat a spontaneous or induced arrhythmia (List separately in addition to code for primary procedure) (Use 93655 in conjunction with 93653, 93654, 93656)
93656	Comprehensive electrophysiologic evaluation including transseptal catheterizations, insertion and repositioning of multiple electrode catheters with induction or attempted induction of an arrhythmia with atrial recording and pacing, when possible, right ventricular pacing and recording, His bundle recording with intracardiac catheter ablation of arrhythmogenic focus, with treatment of atrial fibrillation by ablation by pulmonary vein isolation
93657	Additional linear or focal intracardiac catheter ablation of the left or right atrium for treatment of atrial fibrillation remaining after completion of pulmonary vein isolation (List separately in addition to code for primary procedure)
93799	Unlisted cardiovascular service or procedure
HCPCS Level II Codes	
None	

Note: Inclusion on this list does not guarantee coverage

NOTE: Guidance Development Framework will be added later

DRAFT

EbGS Coverage Guidance Nuclear Cardiac Imaging

Oregon Health Evidence Review Commission
Evidence-based Guidelines Subcommittee
June 5, 2014

Oregon Health Authority

Oregon Health & Science University

Center for Evidence-based Policy

Background

Multiple testing options to diagnose CAD

- Angiography is gold standard
  - Invasive
  - Does not assess perfusion
- Non-invasive alternatives
  - ETT (exercise treadmill test)
  - ECHO (echocardiogram)
  - SPECT (single photon emission computed tomography)
  - PET (positron emission tomography)
  - Other imaging options (CCTA, CACS, cardiac MRI)

Evidence Review

Source report: Washington State Health Care Authority Health Technology Assessment Program. (2013). Cardiac Nuclear Imaging. Olympia, WA: Health Technology Assessment Program. Retrieved December 2, 2013, from http://www.hca.wa.gov/hta/Pages/nuclear.aspx

Evidence Review

Effectiveness evaluated by population

- Asymptomatic at high risk* of CAD
- Symptomatic at low-to-intermediate risk* of CAD
- Symptomatic at high risk* of CAD
- Known CAD
- Mixed populations

Other outcomes assessed:

- Diagnostic accuracy
- Extra-cardiac findings
- Risks of testing
- Comparative value

*Risk of CAD defined using Diamond-Forrester model or similar tool:

Probability of CAD < 10% = Low
Probability of CAD 10% to 90% = Intermediate
Probability of CAD > 90% = High

Evidence Review - Effectiveness

Asymptomatic at High Risk of CAD

- No difference in mortality or CV events compared to no screening over 5 years
- Increased short-term rates of referral for angiography and revascularization with SPECT

Symptomatic at Low-to-Intermediate Risk CAD

- SPECT vs. ETT – no difference in mortality, MACE, revascularization, QoL
- Stress SPECT vs. stress ECHO – no difference in ability to predict MACE (both better than ETT or rest ECHO)

Evidence Review - Effectiveness

Symptomatic at High Risk of CAD

- SPECT vs. ETT – higher rate of revascularization, referral for further testing with ETT; no evidence on mortality or MACE
- SPECT vs. PET – no difference in revascularization, higher rate of referral for angiography in PET
- Stress SPECT vs. stress ECHO – no difference in event-free survival at 7 years

Known CAD

- SPECT post + angiography resulted in lower rate of revascularization than no SPECT (36% vs. 53%)
- PET altered patient management in 78% of patients

Evidence Review - Effectiveness

Mixed Populations

- Stress SPECT vs. stress ECHO vs. stress cMRI vs. angiography- no difference in mortality, QoL; more cardiac events with stress ECHO (1 patient with multiple hospital admissions; UK study)
- SPECT vs. PET – no difference in mortality or MI; lower rates of CABG, any revascularization, referral for angiography with PET
- SPECT vs. ECHO – no difference in event rates, event-free survival
- SPECT vs. ECHO vs. cMRI – no difference in referral for angiography

Evidence Review – Diagnostic Accuracy

- Stress SPECT vs. stress ECHO (meta-analysis of 44 studies) – no difference in sensitivity (85% to 87%) or specificity (~ 77%)
- Three other meta-analyses produce slightly different numbers

Evidence Review – Extra-cardiac Findings

Incidence of extra-cardiac findings very low for SPECT (0 to 0.3%), much lower than CCTA (7% to 69%)

Evidence Review – Risks of Testing

- Stress SPECT vs. ETT – no difference in adverse events
- Most adverse events related to stressor agent
- Radiation exposure
  - SPECT 7-30 mSv
  - PET 2-14 mSv
  - CCTA 2-14 mSv
  - Angiography 5-7 mSv
  - ETT 0 mSv
  - ECHO 0 mSv

Evidence Review – Comparative Value

- SPECT vs. ETT –
  - US study: costs higher for SPECT ($643) due to higher initial cost; outweighs increased costs of more downstream testing in ETT ($338)
  - UK study: costs higher for ETT ($1,244) due to increased costs of more downstream testing; SPECT cost $743
- Stress SPECT vs. stress ECHO vs. stress cMRI vs. angiography (UK study)
  - angiography lowest cost
  - Incremental cost similar for SPECT and cMRI similar (~$650)
  - Highest cost for ECHO (~$1,250)
- SPECT vs. PET – PET may be less costly (2 studies)

Evidence Review – Comparative Value

Decision-analytic model developed by WA HTA authors for 7 management approaches, combining costs with evidence:

- Asymptomatic, high-risk individuals
  - SPECT vs. no screening – high certainty of a comparable net health benefit, low value
  - SPECT vs. ETT or ECHO – insufficient evidence
  - PET vs. any alternative – insufficient evidence
- Symptomatic individuals at low-to-intermediate CAD risk
  - SPECT vs. ETT – moderate certainty of a comparable net health benefit, low value
  - SPECT vs. ECHO – high certainty of a comparable net health benefit, reasonable/comparable value
  - PET vs. any alternative – insufficient evidence
- Symptomatic individuals at high CAD risk
  - SPECT vs. ETT – moderate certainty of a small net health benefit, reasonable/comparable value
  - SPECT vs. ECHO – high certainty of a comparable net health benefit, comparable/reasonable value
  - PET vs. any alternative – insufficient evidence
- Known CAD
  - SPECT vs. ETT – insufficient evidence
  - SPECT vs. ECHO – high certainty of a comparable net health benefit, comparable/reasonable value
  - PET vs. any alternative – insufficient evidence

Evidence Review – Comparative Value

Questions?

HEALTH EVIDENCE REVIEW COMMISSION (HERC)

COVERAGE GUIDANCE: NUCLEAR CARDIAC IMAGING

DRAFT for EbGS Meeting Materials 6/5/2014

HERC COVERAGE GUIDANCE

PET is not recommended for coverage for screening or diagnosis of coronary artery disease (CAD) *(strong recommendation).*

Option 1:

Single photon emission computed tomography (SPECT) is not recommended for coverage for screening or diagnosis of CAD *(strong recommendation)* **including** those who are symptomatic with a high risk of CAD *(weak recommendation).*

Option 2:

SPECT is not recommended for coverage for screening or diagnosis of CAD *(strong recommendation)* **except** those who are symptomatic with a high risk of CAD and have an indeterminate exercise treadmill test (ETT) *(weak recommendation).* In the symptomatic high-risk group, SPECT is superior to ETT but equivalent to stress echocardiogram (ECHO), and with the additional potential harm of radiation.

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

RATIONALE FOR GUIDANCE DEVELOPMENT

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

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

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

DRAFT

EVIDENCE SOURCE

Washington State Health Care Authority Health Technology Assessment Program. (2013). *Cardiac Nuclear Imaging.* Olympia, WA: Health Technology Assessment Program. Retrieved December 2, 2013, from http://www.hca.wa.gov/hta/Pages/nuclear.aspx

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

SUMMARY OF EVIDENCE

Clinical Background

Coronary artery disease (CAD) is among the most common chronic conditions in the U.S., affecting over 16 million adults. Due to its prevalence, and because several options (e.g., surgery, medication) exist to reduce CAD-related morbidity and mortality, accurate diagnosis and/or risk stratification of CAD is critical. Currently the definitive standard for diagnosis is invasive coronary angiography. Because angiography primarily documents the anatomic presence of significant stenosis rather than identifying the "culprit" lesions likely to cause an adverse cardiovascular event, a growing number of non-invasive tests have been developed to identify CAD lesions significant enough to affect the flow of blood to the heart (i.e., myocardial perfusion). These functional tests are typically performed under exercise- or pharmacologically induced stress to determine whether blood flow deteriorates when the stressor is introduced.

The most common tests of cardiac function include the stress-electrocardiogram (ECG), or treadmill test (ETT), which measures cardiac activity via electrical signals, and the echocardiogram (ECHO), which uses ultrasound to measure abnormalities in heart wall motion using 2-dimensional imagery. ETT has fallen out of favor for use in patients at higher risk of CAD, however, as it has relatively low sensitivity in these patients, while stress-ECHO has been found to lack precision in detecting single-vessel versus multi-vessel disease and may produce suboptimal imagery in obese patients, those with chronic respiratory conditions, and patients with chest deformities or pre-existing myocardial damage.

To address some of these concerns, "nuclear imaging tests" have been developed to provide perfusion data in a broader spectrum of patients. The most longstanding of these is single photon emission computed tomography (SPECT), which uses a radioactive tracer and gamma camera to obtain 3-dimensional images of tracer uptake; areas of poor uptake are associated with abnormal levels of perfusion. Positron emission tomography (PET) scanners are also used with a radiotracer, and are felt by some to provide better image resolution in heavier patients and those with dense breast

tissue. So-called “hybrid” modalities have also been introduced to visualize both perfusion abnormalities and anatomic lesions using CT or MRI imagery in addition to nuclear testing.

There are trends in the use of cardiac nuclear imaging tests that are currently points of controversy, however. For one, the use of nuclear imaging for cardiovascular testing has grown substantially in recent years. In addition, questions have been raised about the appropriateness of nuclear imaging in certain populations. A substantial decrease in the prevalence of abnormal findings on such tests has been observed over time, due in part to greater recognition and treatment of cardiac risk factors but also to possible changes in referral patterns. This combination of substantial growth in utilization of cardiac nuclear imaging and declining rates of “positive” test results raises questions about the populations and indications for which such testing is appropriate.

Evidence Review

In the Washington HTA report, “symptomatic” means a patient with symptoms suggestive of myocardial ischemia (symptoms not specified). Risk categories of low, moderate and high were defined by the authors based on the Diamond-Forrester model of pretest probability, which incorporates age, gender and type of chest pain. These categories equate to probability ranges of <10%, 10-90% and >90% respectively. However, when other risk classification systems were used in the included studies, that information was utilized and reported by the authors.

Comparative Clinical Effectiveness

Asymptomatic Patients at High Risk of CAD

The one available study assessing the impact of cardiac nuclear imaging in asymptomatic, high-risk patients found no difference between SPECT screening and no screening in mortality or cardiovascular events, although many patients in both groups received subsequent stress testing for clinical reasons over approximately 5 years of follow-up. SPECT screening did increase the short-term rates of referral for angiography and revascularization vs. no screening.

Symptomatic Patients at Low-to-Intermediate Risk of CAD

Correlation of Imaging Study Findings with Mortality and Cardiovascular Events
Rates of mortality and major cardiac adverse events (MACE) did not generally differ between imaging modalities in available studies. Patients in the WOMEN study, an RCT of 772 women randomized to SPECT or ETT-based testing strategies were at very low CAD risk. The rates of all major adverse cardiovascular events at 2 years were 1.7% and 2.3% for ETT and rest/stress SPECT respectively, but this difference was not

significant (Hazard Ratio [HR]: 1.3; 95% CI: 0.5, 3.5; p=.59). The rate of revascularization also did not statistically differ between groups.

The long-term prognostic value of exercise SPECT, exercise ECHO, ETT, and clinical parameters was measured in a single cohort of 248 patients who were followed for a mean of 3.7 years. A total of 64 MACE occurred during follow-up. In multivariate analyses examining the incremental impact of (1) clinical + ETT data; (2) data in (1) + rest ECHO data; (3) data in (1) + exercise ECHO data; and (4) data in (1) + exercise SPECT data on predicting MACE events, the area under the curve[1] did not statistically differ between the SPECT and ECHO models (0.78 and 0.77 respectively), but was significantly (p<.05) higher than the base model (0.68) or the rest ECHO model (0.72).

One study evaluated the impact on all-cause mortality of normal findings on stress-only vs. stress/rest SPECT protocols in nearly 17,000 low-to-intermediate risk patients followed for a median of 4.5 years. Annualized unadjusted mortality rates were statistically-significantly greater in the stress/rest group (2.92% vs. 2.57% for stress-only, p=.02); however, this difference was no longer apparent after multivariate adjustment for differences in baseline characteristics. The authors conclude that a stress/rest protocol may be unnecessary in lower-risk individuals.

Downstream Testing and Clinical Decision-Making

The impact of testing on downstream resource utilization and clinical decisions was evaluated only in the WOMEN study. Over 2 years of follow-up, repeat testing with the same modality was more frequent in the SPECT group vs. ETT (9% vs. 3%), although this difference was not statistically tested. However, 18% of women randomized to ETT crossed over to SPECT during follow-up. The overall rate of referral to angiography was higher in the ETT group (9.0% vs. 5.5% for SPECT, p<.0001). Changes in the use of nitrates, beta-blockers, and antidepressant therapies during follow-up did not differ between the two arms in the study.

Health-related Quality of Life

The impact of testing on health-related quality of life (HrQOL) also was examined only in the WOMEN study. Similar proportions of women in each treatment group reported “excellent” or “very good” QoL as well as “best” or “average” life satisfaction, with no statistical differences between groups. There were also no statistically-significant differences between ETT and SPECT groups in relation to changes in any of the subscales.

[1] This measure refers to the receiver operating characteristic (ROC) curve, a figure depicting the power of a diagnostic test. It includes both test sensitivity and specificity. A ROC curve for a perfect test has an area under the ROC curve of 1.0, while a test that performs no better than chance has an area under the curve of 0.5.

Symptomatic Patients at High Risk of CAD

Correlation of Imaging Study Findings with Mortality and Cardiovascular Events
In high risk populations, some differences in event rates by modality were apparent. An RCT of ETT vs. SPECT in 457 intermediate-to-high risk patients focused primarily on the period between testing and diagnosis, but did report on the rate of revascularization, which occurred more frequently in the ETT group (18% vs. 11% for SPECT, not statistically tested). In the "SPARC" registry, a study comparing short-term outcomes of PET, SPECT and coronary CT angiography (CCTA), revascularization rates at 90 days did not materially differ between PET and SPECT, regardless of whether findings were mildly or moderately-severely abnormal. Neither of these studies evaluated longer term outcomes such as mortality or cardiovascular events.

Another study assessed the prognostic value of both dobutamine ECHO and dobutamine SPECT in 301 patients who were unable to exercise and were at intermediate-to-high risk of CAD; patients were followed for a mean of 7.3 years. Event-free survival was significantly better for patients with normal vs. abnormal findings on both tests, and did not differ statistically between tests. In multivariate models based on clinical data, stress testing, and imaging results, abnormal findings on either SPECT or ECHO were the strongest predictors of both cardiac death (HR [95% CI]: 4.4 [1.2, 21.0] and 3.4 [1.2, 12.0] for SPECT and ECHO respectively) and cardiac events (3.1 [1.1, 8.9] and 2.6 [1.1-6.2] respectively).

Downstream Testing and Clinical Decision-Making
Two studies reported on the effects of testing on downstream resource use and/or clinical decisions. Of the 207 patients randomized to ETT in one RCT, a total of 146 (71%) were referred for further testing (47% to angiography and 23% to stress ECHO). In contrast, further testing was requested in only 16% of patients randomized to SPECT, all of which were angiography procedures ($p<.0001$ for the comparison). ETT also appeared to generate more false-positives for significant CAD. Only 38% of ETT patients referred to angiography were revascularized, vs. 66% of SPECT patients so referred ($p<.05$).

In a registry study, referral for angiography occurred in a greater percentage of PET patients (11.1% vs. 4.3% for SPECT; $p<.001$). In multivariate analyses controlling for patient characteristics, comorbidities, and testing location, imaging modality was significantly and positively correlated with referral to angiography for PET (OR: 5.0; 95% CI: 1.0, 24.4) in comparison to SPECT. Neither PET nor SPECT were associated with significant medication changes.

Health-Related Quality of Life
There were no studies in symptomatic, high-risk individuals that reported on the impact of cardiac nuclear imaging tests on HrQoL.

Known CAD

Correlation of Imaging Study Results with Mortality and Cardiovascular Events
One comparative cohort study compared the rate of revascularization in 2,951 patients with known CAD and left ventricular dysfunction and (1) who had been tested with SPECT before referral for angiography; (2) were tested with SPECT only after a positive angiography; or (3) had no SPECT before or after angiography. The rate of revascularization differed significantly (p=.001) among groups, with the lowest rate of 35.8% seen in postangiography SPECT patients, 45.6% in patients who had SPECT pre-angiography, and 53.2% among patients undergoing angiography with SPECT neither before nor afterward.

Downstream Testing and Clinical Decision-Making
In one study, 100 consecutive patients, 79% of whom had known CAD, underwent rest-stress PET perfusion testing. Physicians were first queried on proposed patient management strategies without PET perfusion data; actual patient management was measured 4 weeks after PET. Proposed patient management was altered in 78% of patients. Most prominently, conservative medical management was initially proposed in 28% of patients; after PET testing, 76% were managed this way in actuality. In addition, use of angiography to guide treatment via PTCA was proposed in 6%, but was performed in 20% after PET testing.

Health-Related Quality of Life
There were no studies in patients with known CAD that reported on the impact of cardiac nuclear imaging tests on HrQoL.

No comparative studies evaluating the impact of serial nuclear imaging in asymptomatic patients with known CAD were identified.

Mixed Populations

The largest number of studies was available for populations that did not fit neatly into the categories described above. They represented a true “mix” of patients based on relatively uniform distributions by risk or pretest probability, presence or absence of symptoms, and/or inclusion of patients with known vs. suspected CAD. A total of 10 studies were identified.

Correlation of Imaging Study Results with Mortality and Cardiovascular Events
Data on mortality and cardiovascular events were available in 8 studies. The Cost-Effectiveness of Functional Cardiac Testing (CeCAT) Trial was an RCT comparing multiple diagnostic strategies—rest-adenosine stress SPECT, ECHO (dobutamine stress), adenosine stress cardiac magnetic resonance imaging (cardiac MR), and direct referral to angiography—among 898 primarily high-risk patients with known or suspected CAD and stable symptoms of ischemia who were referred to a tertiary center

in the UK for angiography and were followed for 18 months. In this study, the number of total, cardiac, and noncardiac deaths did not statistically differ by imaging modality. When compared with the referent angiography group, the number of nonfatal adverse cardiac *events* did not differ for SPECT or cardiac MR, but was statistically-significantly higher for ECHO (relative risk [RR]: 1.95; 95% CI: 1.23, 3.08; p=.012), primarily because of more admissions for chest pain. When the number of *patients* reporting adverse cardiac events was compared, however, no significant differences were observed (one patient in the ECHO group was responsible for seven hospital admissions).

Findings from a study comparing PET and SPECT were somewhat mixed. No differences in cardiovascular mortality or the rate of MI were observed between groups. However, the rates of CABG (3.4% vs. 7.8%, p<.01) and any revascularization (6.0% vs. 11.4%, p<.01) were statistically-significantly lower for PET vs. the internal (identified by report authors) SPECT control group. The rate of any revascularization was also significantly lower in comparison to the external (using results from another published trial) SPECT control group (6.0% vs. 13.0%, p<.0001).

Three cohort studies comparing the prognostic ability of SPECT and ECHO generally showed comparable results for both tests. No statistical differences between imaging modalities in event rates or event-free survival were observed in 2 studies. In the third, an evaluation of exercise stress ECHO vs. exercise stress SPECT in 206 symptomatic veterans who received both tests and were followed for up to 10 years, moderate-to-large ischemia on ECHO was the strongest independent predictor of overall mortality (RR: 6.2; p<.0001), cardiovascular death (RR: 17.6; p=.01), congestive heart failure (RR: 17.4; p=.0005), or sudden death (RR: 26.8; p=.003). The presence of moderate-to-large fixed defects on SPECT was the strongest independent predictor of nonfatal MI (RR: 8.1; p=.0002) and unstable angina (RR: 3.0; p=.005).

One study assessed the predictive capability of functional data from ETT, exercise stress SPECT, and the “Gensini score” from angiography evaluation in 732 patients who were followed for a mean of 3.5 years. Abnormal results on SPECT and the Gensini score were significantly (p≤.01) associated with poorer event-free survival, while ETT data were not. Analyses of the receiver operator curve (ROC) for events indicated that SPECT was the strongest independent predictor of events (0.67 vs. 0.61 and 0.46 for Gensini score and ETT, p<.05).

Downstream Testing and Clinical Decision-Making

A total of three studies examined the impact of cardiac nuclear imaging on further testing and clinical decision-making. In the CeCAT trial, the proportions of patients in the SPECT, ECHO, and cardiac MR groups who were referred to angiography ranged between 75-80% and did not statistically differ between groups; in addition, decisions on

further invasive or medical management were also similar. The rate of referral to angiography in the study comparing PET and SPECT was statistically-significantly lower for PET (13%) in comparison to both the internal (identified by report authors) and external (using results from another published trial) SPECT groups (31% and 34% respectively, p<.0001). The rate of angiography negative results was also significantly lower for PET vs. internal SPECT controls (5.2% vs. 15.6%, p<.0001).

Finally, a hypothetical referral rate to angiography was assessed in 955 patients undergoing ETT and rest-exercise stress SPECT. Algorithms using ETT data alone, SPECT data alone, and a combination of the 2 tests were applied. An estimated 27% of patients would have been referred to angiography based on ETT results alone, vs. 13% for SPECT data alone and 12% using both ETT and SPECT data (p<.01 for both comparisons to ETT alone). Findings were similar when compared among patients without known CAD.

Health-Related Quality of Life
HrQoL was assessed in the CeCAT trial. While some statistically-significant differences were noted in certain subscales at particular time points, improvements in HrQoL were clinically comparable across testing groups for all measures.

Diagnostic Accuracy

A total of 8 studies were available that examined the accuracy of cardiac nuclear imaging tests in relation to a functional reference standard. This is currently believed to be a more accurate method to determine whether a defect noted on non-invasive imaging relates to CAD that is functionally-significant—that is, likely to be the cause of an adverse cardiovascular event if not treated. Recent research has raised questions about the use of anatomic data on angiography to confirm findings of functional tests such as ECHO, SPECT, and PET. There is nevertheless a large body of evidence evaluating the accuracy of noninvasive functional tests using visualization of coronary arteries as the reference standard.

One of the most widely-cited meta-analyses compared the diagnostic accuracy of exercise ECHO and exercise SPECT based on 44 studies. Pooled sensitivity of the 2 tests was similar (85% and 87% for ECHO and SPECT respectively), but pooled specificity was rated higher for ECHO (77% vs. 64% for SPECT, p<.05). However, substantial heterogeneity in study populations, imaging protocols, and SPECT radiotracers was noted for this sample; subsequent reanalysis with controls for heterogeneity found no statistical differences between the tests.

Methods to assess diagnostic accuracy have also evolved, and feature newer techniques designed to capture the natural correlation between sensitivity and specificity. A recent meta-analysis using newer bivariate methods found that ECHO was

slightly more sensitive than SPECT (87% vs. 83% respectively), while SPECT was somewhat more specific (77% vs. 72% for ECHO). An additional bivariate meta-analysis using a much larger set of 113 SPECT studies found greater sensitivity (88%) and similar specificity (76%), although other commentators have noted that the older SPECT studies included in this review were subject to “verification bias” (i.e., use of the reference standard only in test-positive or other selected individuals), which tends to inflate sensitivity and may also reduce specificity. This meta-analysis also included estimates of diagnostic accuracy from 9 PET studies (pooled estimates of 93% and 81% for sensitivity and specificity respectively).

Finally, a third recent meta-analysis estimated diagnostic performance from 114 SPECT and 15 PET studies. SPECT sensitivity was similar to that reported elsewhere (88%), but specificity was somewhat lower (61%). Sensitivity and specificity for PET was estimated to be 84% and 81% respectively.

Other Outcomes

Extracardiac Findings

With the enhanced imagery available for many noninvasive tests, incidental findings outside of the area of interest can be problematic given the additional resources required for investigation. The reported rate of incidental extracardiac findings is very low with nuclear imaging tests given the limited field of detection, however; most available studies are limited to case reports of mediastinal masses. One recent study compared the rate of such findings between CCTA and SPECT in 479 patients; extracardiac findings requiring further investigation were detected in 7% of CCTA patients but in no SPECT patients (p=.0001). Another analysis examined images of 2,155 patients undergoing SPECT studies, 6 (0.3%) of whom had extracardiac findings requiring follow-up. Four of the 6 patients had malignancies requiring further treatment. No PET studies reported on extracardiac findings.

While SPECT itself is associated with a low rate of extracardiac findings, the increasing use of CT for attenuation correction may result in increased detection of these findings. In a cohort study assessing prevalence of extracardiac findings from 582 SPECT/CT studies, a total of 400 (68.7%) included noncardiac findings, 196 (33.7%) of which were felt to be potentially relevant.

Equivocal/Indeterminate Results

While equivocal or indeterminate findings are possible with any diagnostic test, these results are rarely published. A recent systematic review of nearly 1,200 diagnostic accuracy studies found that 35% reported the presence of inconclusive results. Inconclusive results were reported in only one of the studies in this report. In the CeCAT

trial comparing SPECT with ECHO, cardiac MR, and angiography, rates of equivocal findings were 4.0%, 6.6%, 6.6% and 2.0% respectively.

Risks of Testing

Patients appear to be at minimal immediate risk from cardiac nuclear imaging tests in and of themselves, although harms data are reported in only a small number of comparative studies. The risks that are reported are related primarily to the stressor employed (i.e., exercise or pharmacologic stress).

Comparative Data on Testing Risks

Only 2 studies compared adverse effects of multiple testing modalities. In the WOMEN study that randomized patients to ETT or exercise SPECT, no statistically significant differences between groups were noted in rates of chest pain, dyspnea, or fatigue after testing. In the CeCAT trial comparing SPECT, ECHO, cardiac MR, and angiography, specific reasons for failed tests were recorded. Failure to complete the test due to adverse effects occurred in 4 ECHO patients (1.8%), due to vasovagal reactions, blood pressure changes and dyspnea; no patient failed to complete SPECT due to adverse effects.

Adverse Effects by Stressor

Information on adverse effects attributed to specific stressors was obtained from 15 studies. Regardless of the comparisons made, events were typically described as non-serious and resolved once the stressor infusion ended. Reported ranges of adverse effects were similar across pharmacologic agents. Limited data suggest lower rates of adverse effects for exercise vs. pharmacologic stress in the 2 studies making this comparison, although statistical comparisons were not available for all event types.

Radiation Exposure

Potential adverse health effects associated with radiation exposure are important factors to consider in the evaluation of cardiac nuclear imaging tests, particularly because patients may already be exposed to radiation at other points along the diagnostic pathway (e.g., CCTA, angiography), cumulative radiation dose may be substantial in patients receiving serial imaging studies, and imaging alternatives such as ECHO and cardiac MR exist that do not involve radiation. Radiation dose is a measure of ionizing energy absorbed per unit of mass, expressed as units of Gy (Gray) or mGy; it often is quoted as an equivalent “effective” dose to major organs in the scanned area, in units of Sv (Sievert) or mSv. For x-rays, the radiation type produced by CT scanners, 1 mSv = 1 mGy. Average total effective dosages for SPECT range from 7 to 30 mSv, while for PET and CCTA the range is 2 to 14 mSv, and for invasive coronary angiography the range for is 5 to 7. While exposure to ionizing radiation at these levels

is associated with potential increase in cancer risk, the latency period for the development of such cancers may range from 10 to 40 years for solid tumors depending on the age and sex of the patient being tested. The intended use of cardiac imaging tests then becomes a critical consideration.

Differential Effectiveness/Safety for Key Patient Subgroups

The *comparative* impact of cardiac nuclear imaging tests vs. alternative testing strategies in certain subgroups is presented below.

Patient Age, Sex, Race or Ethnicity, and Comorbidities

A single comparative cohort study was available that assessed all-cause mortality for stress only vs. stress-rest SPECT (n=16,854) in specific subgroups over a mean of 4.5 years of follow-up. On a univariate basis, stress-rest protocols were associated with a statistically-significantly higher mortality rate in older (age >65) individuals, men, patients with a BMI <30 kg/m2, and patients with diabetes. However, after multivariate adjustment for baseline characteristics, no statistically-significant differences remained. Several large cohort studies and meta-analyses have assessed the performance of SPECT in certain patient subgroups. For example, several studies have found that SPECT's diagnostic and prognostic performance is similar for women and men. Comparable results have also been found in several large ECHO studies. A meta-analysis of risk-stratification studies in over 13,000 patients age >65 years found that both stress SPECT and stress ECHO accurately risk-stratified patients vs. ETT. A multicenter cohort study of approximately 1,100 patients found that SPECT results were predictive of cardiac events in both Caucasian and African-American patients.

Analyses comparing patients with and without diabetes suggest that, while diabetes is a predictor of mortality for any nuclear imaging result, SPECT testing provides incremental prognostic information in patients with and without diabetes alike. Multiple studies have found that SPECT is feasible and has comparable diagnostic and prognostic performance in normal-weight, overweight, and obese patients. Finally, a meta-analysis of SPECT and ECHO studies in hypertensive patients showed diagnostic accuracy similar to that observed in all patients with suspicion of CAD.

Clinical Setting

In a comparison of stress-only vs. stress-rest SPECT, mortality was initially statistically-significantly higher in stress-rest patients in an inpatient setting. After multivariate adjustment, however, no significant differences remained. Limited additional data are available explicitly comparing the performance of SPECT by setting. One study evaluating the potential benefit of an emergency department chest pain clinic estimated that unnecessary hospitalizations would be reduced in 30% of patients and inappropriate discharges avoided in 6% through the use of a selective SPECT protocol.

Selection of Test by Primary Care vs. Specialty Physician

No study assessed the impact of ordering specialty on patient outcomes, clinical decision-making, or costs. There are, however, several studies that have assessed the impact of specialty on whether ordered cardiac SPECT studies meet published appropriate use criteria (AUC). In a multicenter assessment of an online SPECT appropriateness classification system, one study found that the rate of inappropriate studies was statistically-significantly higher among non-cardiologists (19.5% vs. 13.2% for cardiologists, p<.0001). Similar findings have been observed in several single center studies. Of note, most inappropriate ordering of SPECT perfusion studies appears to have occurred in women, younger patients, and/or those without symptoms.

Scan Vendor, Type of Assessment, Type of Radioisotope, and Type of Stressor

No study assessed the impact of scan vendor or qualitative vs. quantitative assessment on patient outcomes, clinical decision-making, or costs. Most of the studies evaluating differences according to stressor type focused on rates of adverse effects of pharmacologic testing. The study that evaluated stress-only vs. stress-rest SPECT found no statistically-significant effects on mortality with subgroups defined by exercise vs. pharmacologic stress.

Two studies examined the impact of different SPECT radiotracers on outcomes. In one, a total of 1,818 patients underwent exercise or pharmacologic stress SPECT with Tc-99m sestamibi or Tc-99m tetrofosmin. Patients were followed for a mean of 1.5 years, during which no statistically-significant differences were observed between groups in the rates of overall mortality, cardiovascular mortality, or the composite endpoint of cardiovascular mortality or nonfatal MI.

The other study compared mortality outcomes among 2,147 patients with known CAD undergoing pharmacologic stress SPECT with either Tc-99m sestamibi or Tc99m tetrofosmin who were followed for a median of 4 years. During follow-up, a total of 704 all-cause deaths (493 cardiovascular-related) were reported. There was no significant difference in either overall or cardiovascular mortality between radiotracer groups on both an unadjusted and multivariate-adjusted basis.

Analysis of Comparative Value

Limited evidence is available that directly measured and compared the economic impact of non-invasive testing strategies for CAD. Three RCTs compared costs of SPECT to other imaging. In the only economic study performed in the US, an RCT of ETT vs. SPECT in 772 women at low-to-intermediate risk of CAD in 43 cardiology practices across the U.S., total mean costs of testing over 2 years were higher in the SPECT arm ($643 vs. $338, p<.001), as the higher costs of initial SPECT testing outweighed the increased costs of downstream testing in the ETT arm. In another 2-year RCT

conducted in 457 primarily intermediate-risk patients in the UK, however, downstream testing costs were substantially higher in the ETT arm, leading to significantly higher total costs from randomization to diagnosis using National Health Service (NHS) estimates ($1,244 v $743 for SPECT, p<.001). The final UK RCT compared costs of initial and repeat testing, treatment, and adverse events over 18 months of follow-up for mixed-risk patients randomized to SPECT, ECHO, cardiac MR, or direct referral to angiography. Direct referral to angiography was the lowest-cost strategy. Incremental costs (relative to angiography) were similar for the SPECT and cardiac MR strategies (~$650), but were twice as high for patients in the ECHO group (~$1,250) due to a higher rate of hospital readmissions.

Economic evidence for PET was limited to 2 studies. In one, an evaluation of planned vs. actual management before and after PET perfusion testing in 100 patients with known CAD, savings from reduced need for angiography were greater than the incremental costs of PET testing and revascularization, leading to overall savings of $240 per patient. In the other, a matched comparative cohort analysis of PET and SPECT, mean costs of all diagnostic testing were approximately $2,500 in both groups, but greater requirements for revascularization at 1 year led to higher total costs in the SPECT group ($5,937 vs. $4,110 for PET).

Because evidence is limited comparing the short-term clinical consequences and costs for all relevant non-invasive strategies for CAD diagnosis, the authors of this report developed a decision-analytic model to provide additional information. The target population involved men and women with suspected or known CAD who had stable symptoms of myocardial ischemia (i.e., atypical or typical chest pain or other symptoms such as dyspnea). Model outcomes and costs were estimated over a 90-day period. The authors of the Washington HTA report developed 7 different strategies, alone and in combination, to capture a wide range of management approaches:

1. ECHO, followed by invasive coronary angiography if ECHO is positive or inconclusive
2. ETT, followed by angiography if ETT is positive or inconclusive
3. SPECT, followed by angiography if ETT is positive or inconclusive
4. PET, followed by angiography if ETT is positive or inconclusive
5. ETT, followed by ECHO if ETT is positive or inconclusive, followed by angiography if the ECHO is positive or inconclusive
6. ETT, followed by SPECT if ETT is positive or inconclusive, followed by angiography if the SPECT is positive or inconclusive
7. ETT, followed by PET if ETT is positive or inconclusive, followed by angiography if the PET is positive or inconclusive

Because the underlying CAD prevalence varies in different patient populations, the authors calculate results of the identical testing strategies for a population with 10%, 30%, 50% and 70% CAD prevalence. Comparing these results demonstrates the importance of the underlying prevalence on the relative balance of false negatives, false positives, rates of referral to angiography, and costs. For example, among a patient population with a CAD prevalence of 10%, the difference in false negatives between SPECT and ECHO almost vanishes (4 per 1,000). In contrast, the difference in false positives between SPECT and ECHO in a population with 50% CAD prevalence was 33 per 1,000 but is increased to 60 per 1,000 when the underlying prevalence of CAD is only 10%. The relative differences in angiography referral, patients exposed to radiation, and costs also shift.

The authors of this report have devised their own evidence rating system, and reach the following conclusions for specific populations:

- Asymptomatic, high-risk individuals
  - SPECT vs. no screening – high certainty of a comparable net health benefit, low value
  - SPECT vs. ETT or ECHO – insufficient evidence
  - PET vs. any alternative – insufficient evidence
- Symptomatic individuals at low-to-intermediate CAD risk
  - SPECT vs. ETT – moderate certainty of a comparable net health benefit, low value
  - SPECT vs. ECHO – high certainty of a comparable net health benefit, reasonable/comparable value
  - PET vs. any alternative – insufficient evidence
- Symptomatic individuals at high CAD risk
  - SPECT vs. ETT – moderate certainty of a small net health benefit, reasonable/comparable value
  - SPECT vs. ECHO – high certainty of a comparable net health benefit, comparable/reasonable value
  - PET vs. any alternative – insufficient evidence
- Known CAD
  - SPECT vs. ETT – insufficient evidence
  - SPECT vs. ECHO – high certainty of a comparable net health benefit, comparable/reasonable value

DRAFT

- PET vs. any alternative – insufficient evidence

[Evidence Source]

Evidence Summary

In **asymptomatic individuals** at high risk of CAD, there is no evidence of benefit for SPECT screening compared to no screening. In **symptomatic patients at low to intermediate risk of CAD**, evidence is conflicting with regard to ability to predict mortality and cardiovascular events, with one study finding no difference between ETT and SPECT, and another finding that stress SPECT and stress ECHO were better predictors than ETT and rest ECHO. In **symptomatic patients at high risk of CAD**, evidence is conflicting regarding rates of revascularization in those who undergo ETT compared to SPECT. Prognostic value does not differ between stress ECHO and stress SPECT. In **populations with mixed risk of CAD**, stress SPECT, stress ECHO, stress CMR and angiography do not differ in subsequent death or patient reported adverse cardiac events. SPECT and ECHO have similar prognostic abilities, and those tests as well as cardiac MR result in similar proportions of referrals to angiography or change in medical management.

With regard to diagnostic accuracy, SPECT and ECHO have similar sensitivity (83% to 87%) and specificity (64% to 77%), although some analyses suggest that ECHO may be slightly more sensitive and SPECT may be slightly more specific. Extracardiac findings (which may require additional evaluation) are identified rarely with SPECT, and significantly less frequently than CCTA.

Comparative evidence on the risks of various testing strategies is very limited, with the only apparent difference being that exercise stress has lower rates of adverse events than pharmacologic stress. SPECT has the highest radiation exposure of any testing strategy at a range of 7 to 30 mSv. SPECT appears to perform similarly in men and women, Caucasians and African-Americans, normal weight and obese patients, those with and without diabetes and those with and without hypertension.

Evidence is conflicting regarding the value of ETT compared to SPECT. One study suggests that direct referral to angiography is the most cost effective strategy, with SPECT and cardiac MR being of moderate cost, and ECHO being the most costly. Another analysis finds that SPECT has low value compared to no screening in an asymptomatic population and compared to ETT in a low-to-intermediate risk population, and has comparable value compared to ECHO in all other populations.

The evidence pertaining to PET is insufficient to draw conclusions for any outcome.

GRADE-INFORMED FRAMEWORK

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

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variation in values and preferences	Coverage recommendation	Rationale
SPECT in asymptomatic at high risk of CAD	No net benefit, harm of radiation	High	Moderate	Low variability	Not recommended for coverage (strong recommendation)	Framework would lead to IA2a.
SPECT in symptomatic with low/mod risk of CAD	Comparable to ETT and ECHO; greater harms from radiation	Moderate (ETT) High (ECHO)	Moderate	Moderate variability	Not recommended for coverage (strong recommendation)	Framework would lead to IA2a.
SPECT in symptomatic with high risk of CAD	Small health benefit compared to ETT but more harms from radiation Comparable to ECHO but greater harms from radiation	Moderate (ETT) High (ECHO)	Moderate	Moderate variability	*SPECT vs. ETT* Not recommended for coverage (weak recommendation) *SPECT vs. ECHO* Not recommended for coverage (strong recommendation)	*SPECT vs. ETT* Framework would lead to IA1a. However, additional benefit may outweigh harms (upgrading this to a weak for). If ETT is chosen and is indeterminate, SPECT (or ECHO) are reasonable. Echo would be equivalent to SPECT but without harms. *SPECT vs. ECHO* Framework would lead to

DRAFT

Indication/ Intervention	Balance between desirable and undesirable effects	Quality of evidence*	Resource allocation	Variation in values and preferences	Coverage recommendation	Rationale
						IA2a. SPECT is better than ETT but similar to ECHO.
SPECT in known CAD	Unknown compared to ETT Comparable to ECHO but greater harms from radiation	Insufficient (ETT) High (ECHO)	Moderate	Moderate variability	Not recommended for coverage (strong recommendation)	*SPECT vs. ETT* Framework would lead to IIA2. *SPECT vs. ECHO* Framework would lead to IA2a.
PET in all populations	Unknown compared to all interventions	Insufficient	High	Low variability	Not recommended for coverage (strong recommendation)	Framework would lead to IIA2.

*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

There were no quality measures that pertained to this topic identified when searching the National Quality Measures Clearinghouse.

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

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

DRAFT

Appendix A. GRADE Element Descriptions

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

Strong recommendation

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

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

Weak recommendation

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

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

Quality or strength of evidence rating across studies for the treatment/outcome[2]

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

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

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

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

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

DRAFT

Appendix B. Applicable Codes

CODES	DESCRIPTION
ICD-9 Diagnosis Codes	
402.0-402.9	Hypertensive heart disease
411.0	Postmyocardial infarction syndrome
411.1	Intermediate coronary syndrome (impending infarction, preinfarction angina, preinfarction syndrome, unstable angina
411.8	Other acute and subacute forms of ischemic heart disease
411.81	Acute coronary occlusion without myocardial infarction
411.89	Other acute and subacute forms of ischemic heart disease, other
413.0	Angina decubitus (nocturnal angina)
413.1	Prinzmetal angina (variant angina pectoris)
413.9	Other and unspecified angina pectoris (NOS, cardiac, equivalent, of effort, angina syndrome, status anginosus, stenocardia, syncope anginosa)
428.0	Congestive heart failure, unspecified
428.1	Left heart failure
428.20	Systolic heart failure, unspecified
428.21	Acute systolic heart failure
428.22	Chronic systolic heart failure
428.23	Acute on chronic systolic heart failure
428.30	Diastolic heart failure, unspecified
428.31	Acute diastolic heart failure
428.32	Chronic diastolic heart failure
428.33	Acute on chronic diastolic heart failure
428.40	Combined systolic and diastolic heart failure, unspecified
428.41	Acute combined systolic and diastolic heart failure
428.42	Chronic combined systolic and diastolic heart failure
428.43	Acute on chronic combined systolic and diastolic heart failure
428.9	Heart failure, unspecified (cardiac failure, NOS, heart failure NOS, myocardial failure NOS, weak heart
429.2	Cardiovascular disease, unspecified
429.3	Cardiomegaly
429.4	Functional disturbances following cardiac surgery
429.7	Certain sequelae of myocardial infarction not elsewhere classified
429.79	Certain sequelae of myocardial infarction not elsewhere classified, other
429.83	Takotsubo syndrome
429.9	Heart disease, unspecified
786.50	Chest pain, unspecified
ICD-9 Volume 3 (Procedure Codes)	
92.05	Cardiovascular scan and radioisotope function study
CPT Codes	
75557	Cardiac magnetic resonance imaging for morphology and function without

DRAFT

CODES	DESCRIPTION
	contrast material
75559	with stress imaging
75561	Cardiac magnetic resonance imaging for morphology and function without contrast material(s), followed by contrast material(s) and further sequences
75563	with stress imaging
75565	Cardiac magnetic resonance imaging for velocity flow mapping (List separately in addition to code for primary procedure)
75571	Computed tomography, heart, without contrast material, with quantitative evaluation of coronary calcium
75572	Computed tomography, heart, with contrast material, for evaluation of cardiac structure and morphology (including 3D image postprocessing, assessment of cardiac function, and evaluation of venous structures, if performed)
75573	Computed tomography, heart, with contrast material, for evaluation of cardiac structure and morphology in the setting of congenital heart disease (including 3D image postprocessing, assessment of LV cardiac function, RV structure and function and evaluation of venous structures, if performed)
78451	Myocardial perfusion imaging, tomographic (SPECT) (including attenuation correction, qualitative or quantitative wall motion, ejection fraction by first pass or gated technique, additional quantification, when performed); single study, at rest or stress (exercise or pharmacologic)
78452	multiple studies, at rest and/or stress (exercise or pharmacologic) and/or redistribution and/or rest reinjection
78453	Myocardial perfusion imaging, planar (including qualitative or quantitative wall motion, ejection fraction by first pass or gated technique, additional quantification, when performed); single study, at rest or stress (exercise or pharmacologic)
78454	multiple studies, at rest and/or stress (exercise or pharmacologic) and/or redistribution and/or rest reinjection
78459	Myocardial imaging, positron emission tomography (PET), metabolic evaluation
78472	Cardiac blood pool imaging, gated equilibrium; planar, single study at rest or stress (exercise and/or pharmacologic), wall motion study plus ejection fraction, with or without additional quantitative processing
78473	multiple studies, wall motion study plus ejection fraction, at rest and stress (exercise and/or pharmacologic), with or without additional quantification
78481/3	Cardiac blood pool imaging (planar), first pass technique; single study, at rest or with stress (exercise and/or pharmacologic), wall motion study plus ejection fraction, with or without quantification (single/multiple)
78491	Myocardial imaging, positron emission tomography (PET), perfusion; single study at rest or stress
78492	multiple studies at rest and/or stress
78499	Unlisted cardiovascular procedure, diagnostic nuclear medicine

DRAFT

CODES	DESCRIPTION
93000	Electrocardiogram, routine ECG with at least 12 leads; with interpretation and report
93005	tracing only, without interpretation and report
93010	interpretation and report only
93015	Cardiovascular stress test using maximal or submaximal treadmill or bicycle exercise, continuous electrocardiographic monitoring, and/or pharmacological stress; with supervision, interpretation and report
93016	supervision only, without interpretation and report
93017	tracing only, without interpretation and report
93018	interpretation and report only
93350	Echocardiography, transthoracic, real-time with image documentation (2D), includes M-mode recording, when performed, during rest and cardiovascular stress test using treadmill, bicycle exercise and/or pharmacologically induced stress, with interpretation and report
93351	including performance of continuous electrocardiographic monitoring, with supervision by a physician or other qualified health care professional
93454	Catheter placement in coronary arter(s) for coronary angiography, including intraprocedural injection(s) for coronary angiography, imaging supervision and interpretation
93455	with catheter placement(s) in bypass graft(s) (internal mammary, free arterial, venous grafts) including intraprocedural injection(s) for bypass graft angiography
93456	with right heart catheterization
HCPCS Level II Codes	
A9500	Technetium tc-99m sestamibi, diagnostic, per study dose
A9502	Technetium tc-99m tetrofosmin, diagnostic, per study dose
A9505	Thallium tl-201 thallous chloride, diagnostic, per millicurie
A9526	Nitrogen n-13 ammonia, diagnostic, per study dose, up to 40 millicuries
A9555	Rubidium rb-82, diagnostic, per study dose, up to 60 millicuries
A9560	Technetium tc-99m labeled red blood cells, diagnostic, per study dose, up to 30 millicuries
J0280	Injection, aminophyllin, up to 250 mg
J0461	Injection, atropine sulfate, 0.01 mg
J0151	Injection, adenosine for diagnostic use, 1 mg
J1245	Injection, dipyridamole, per 10 mg
J1250	Injection, dobutamine hydrochloride, per 250 mg
J2785	Injection, regadenoson, 0.1 mg

Note: Inclusion on this list does not guarantee coverage

DRAFT

Appendix C. HERC Guidance Development Framework

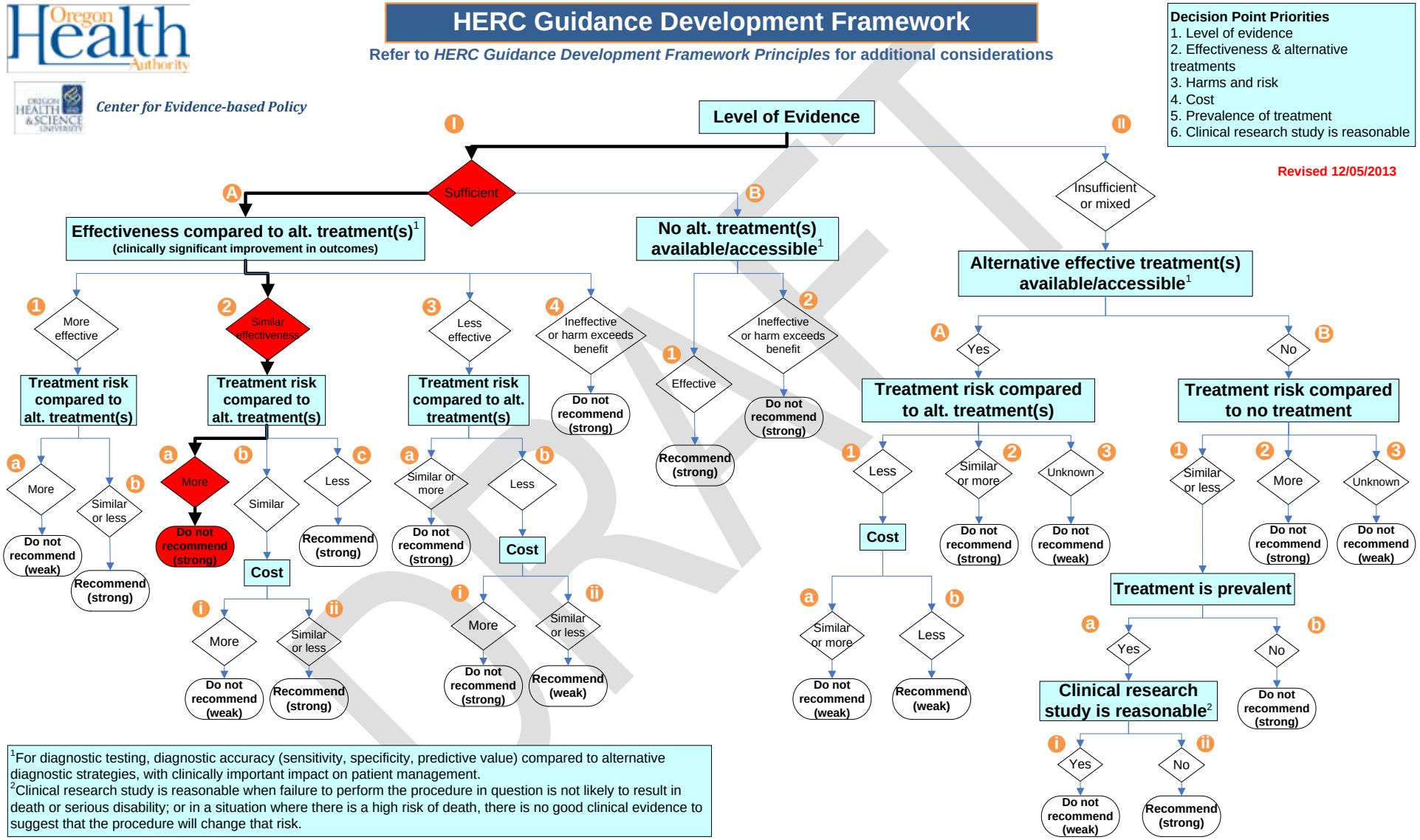
HERC Guidance Development Framework Principles

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

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

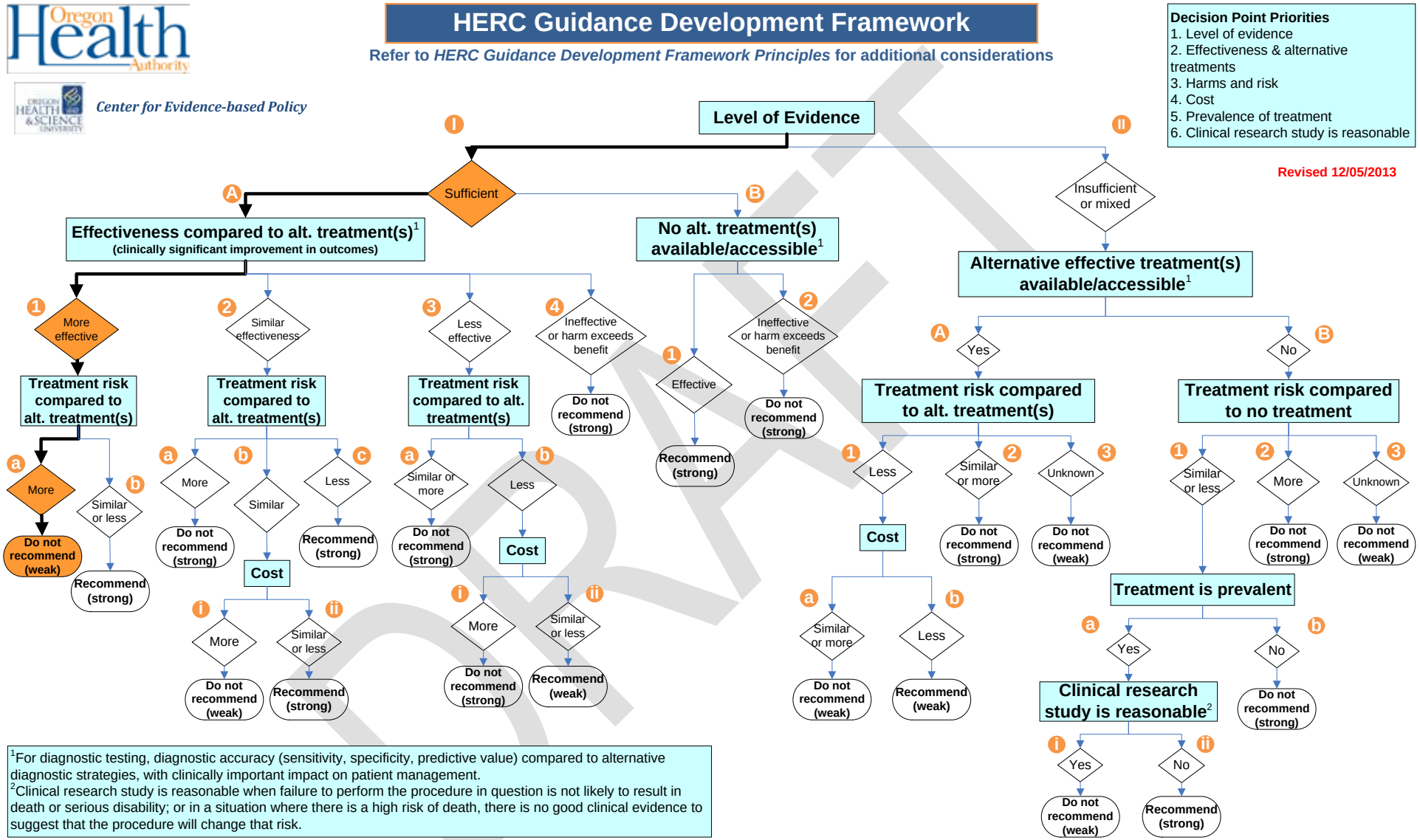
DRAFT

SPECT in: Asymptomatic at high risk of CAD; Symptomatic with low/mod risk of CAD (compared to ETT and ECHO); Symptomatic with high risk of CAD (compared to ECHO); Known CAD (compared to ECHO)

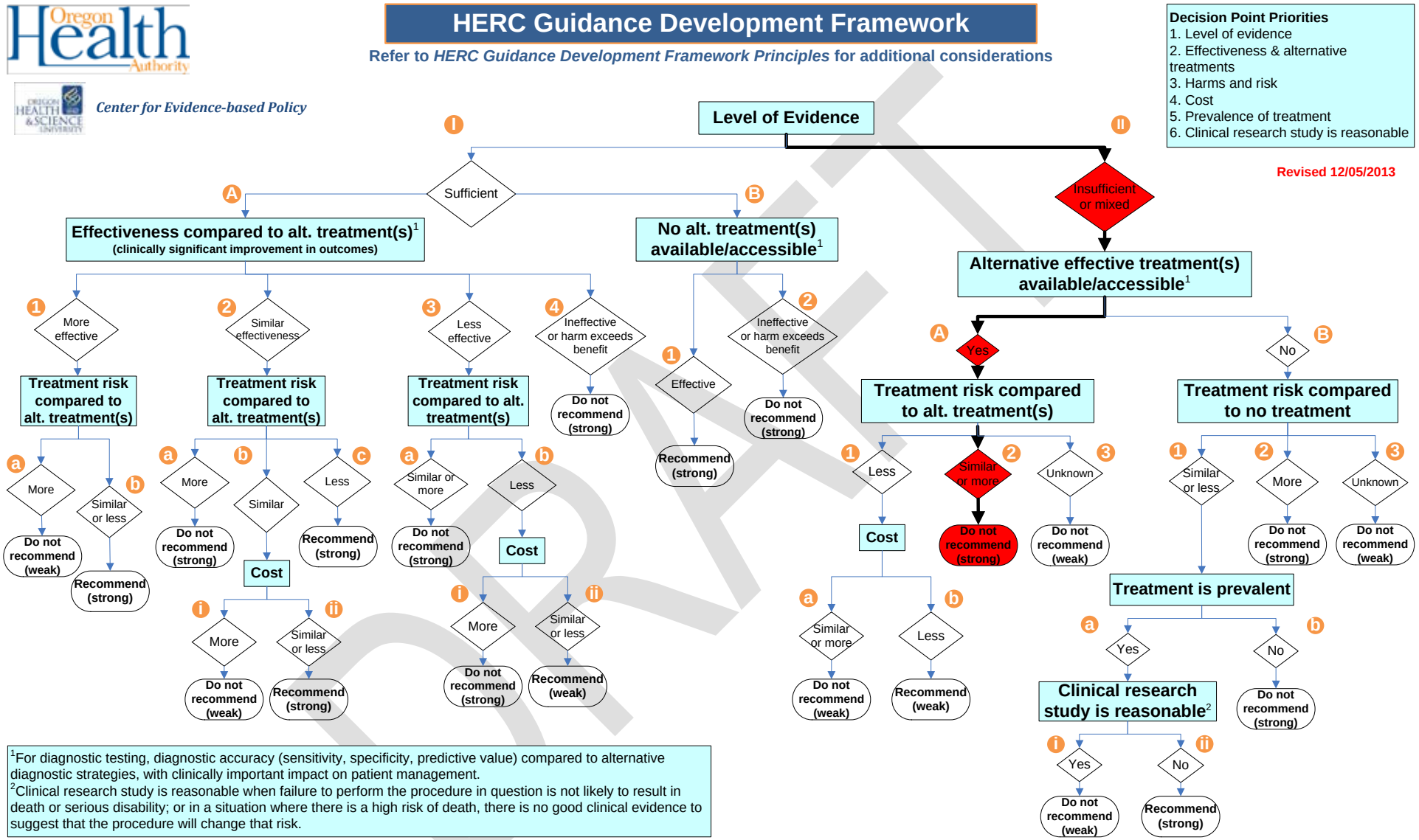


[1]For diagnostic testing, diagnostic accuracy (sensitivity, specificity, predictive value) compared to alternative diagnostic strategies, with clinically important impact on patient management.
[2]Clinical research study is reasonable when failure to perform the procedure in question is not likely to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

SPECT in symptomatic with high risk of CAD (compared to ETT)



SPECT in known CAD (compared to ETT); PET in all populations



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