Comparative Effectiveness Research Review Disposition of Comments Report

Research Review Title: *Noninvasive Technologies for the Diagnosis of Coronary Artery Disease in Women*

Draft review available for public comment from August 30, 2011, to September 26, 2011.

Research Review Citation: Dolor RJ, Patel MR, Melloni C, Chatterjee R, McBroom AJ, Musty M, Wing L, Coeytaux RR, Ross AK, Bastian LA, Anderson M, Kosinski A, Sanders GD. Noninvasive Technologies for the Diagnosis of Coronary Artery Disease in Women. Comparative Effectiveness Review No. 58. (Prepared by the Duke Evidence-based Practice Center under Contract No. 290-2007-10066-I.) AHRQ Publication No. 12-EHC034-EF. Rockville, MD. Agency for Healthcare Research and Quality. June 2012. Available at: www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each comparative effectiveness research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft comparative effectiveness research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Source: http://effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/?pageaction=displayproduct&productID=1019
Published Online: June 12, 2012

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 1	General	Quality of the Report: Superior	Thank you.
Peer Reviewer 1	General	This is an extremely detailed and well written report. The topic is very important clinically as diagnosing women with chest pain syndrome remains a diagnostic challenge. The key questions are clearly stated, important, and appropriate. My primary suggetion is that women with chest pain syndrome, a phrase used throughout the document, is never defined in explicit terms beginning on page 8. In Table 1, page 10, a few examples to assist with a definition are "exertional dyspnea, shortness of breath, and/or angina". This indicates women without chest pain could be included in the group charactericized as having chest pain syndrome. This is problematic and continues to contribute to the problem of how to recognize potential CHD in women who present without chest pain. The label of chest pain syndrome is misleading and indicates all of the women in this group have pain, perpetuating the idea that chest pain is the key indicator of CHD in most/all women. Realizing that many studies use this same terminology, it is understandable that this is used in this review but this needs to be acknowledged and cited as a limitation of this study.	We agree that the definition of chest pain syndrome can be difficult to characterize. We have changed the language to include “suspected CAD” or “symptoms suspicious for CAD,” which includes both angina and nonanginal chest pain and is the term used in the majority of the studies to describe symptomatic women undergoing an evaluation.
Peer Reviewer 1	General	Clarity and Usability: This report is extremely well written, logically organized and the key points are clearly presented. The detailed description of the search strategies and key terms would allow for replication. The conclusions that are identified as having sufficient evidence should be useful for both policy and practice decisions. It is unfortunate that major questions had insufficient evidence but it is one of the strenghts of this report that the authors clearly identify where additional reserach is needed to fully answer these important clinical questions.	Thank you.
Peer Reviewer 2	General	I thought this was a remarkable effort, and mostly successful.	Thank you.
Peer Reviewer 2	General	Overall I found this paper to be readable, careful, comprehensive, and very informative. I expect you will hear other tweaks from invasive cardiologists, but from my "doc in a box" vantage point of benefits and risks for diverse tests the paper was just about right.	Thank you.
Peer Reviewer 2	General	The research plan, and the key questions and comparisons are thoughtful and appropriate.	Thank you.
Peer Reviewer 2	General	I was surprised about how limited in number and size were the studies meeting your eligibility criteria. This certainly highlighted gaps in my mind.	Noted
Peer Reviewer 2	General	I was not convinced by your statement that the missing comparisons could not be made in clinical trials today for ethical reasons. You need to say why.	We are unable to locate this statement in our report and therefore are unable to respond to this comment.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 2	General	I like the focus on sensitivity and specificity and positive and negative predictive values, and NNT.	Thank you.
Peer Reviewer 2	General	Added value is important for screening tests and risk factors, but may not be as relevant for comparative effectiveness.	Noted
Peer Reviewer 2	General	I appreciated the unequivocal declarative conclusions, especially KQ3.	Thank you.
Peer Reviewer 3	General	Quality of the Report: Good	Thank you.
Peer Reviewer 3	General	General Comments: This document is a review of papers published in English over the last 10 years examining diagnostic accuracy, predictors of diagnostic accuracy, risk stratification and prognostic information as well as clinical outcomes of various non-invasive testing modalities in women. The authors included only those studies that reported sex-specific data across the various modalities. The non-invasive modalities included stress electrocardiography, stress echocardiography, myocardial perfusion imaging with SPECT, cardiac MRI and coronary CT angiography encompassing 101 studies and 98 comparative studies in total. They examined the key outcome measures across all studies and separately among "good" studies (quality of the study was determined by application of specific AHRQ criteria).	Noted
Peer Reviewer 3	General	The key questions of this report are appropriate and explicitly stated.	Thank you.
Peer Reviewer 3	General	Clarity and Usability: The report is well structured and organized, and the main points clearly presented. The conclusions drawn from this report can be used to inform policy and practice decisions.	Thank you.
Peer Reviewer 4	General	Thank you for giving me the opportunity to review this manuscript. This is a very challenging topic because, sadly, the data is quite sparse when the gravity of the situation is considered in broad terms. The authors did an excellent job in pooling and evaluating this work. I have no concerns regarding the methodology.	Thank you for reviewing this report.
Peer Reviewer 5	General	Overall an excellent and unbiased resource that will be valuable to inform clinical recommendations and decision making.	Thank you.
Peer Reviewer 6	General	Quality of the Report: Superior	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 6	General	General Comments: I would like to commend the authors for a thorough and timely review of the literature on this topic. This report will be incredibly meaningful in the clinical practice of taking care of women with chest pain syndromes, particularly in light of the choices to make amongst testing modalities and the recent emergence of new modalities. The key questions are clinically meaningful and important for the appropriate diagnostic management of women with chest pain. Currently, the decisions on type of non-invasive test ordered for this population rest mainly in consensus statements from professional societies based upon the knowledge of how each test performs relative to an internal or gold standard, and less on how they compare to one another. As was found in this report there are many gaps in this literature to assist the practitioner and decision-makers in understanding the relative effectiveness of different NITs.	Thank you.
Peer Reviewer 6	General	The authors do not comment on the issue of cost which may play an important role in health care decision-making, especially in the setting of emerging technology. This is reasonable as it is important to focus first on clinical outcomes, but it could be more clearly specified that this is a review of clinical comparative effectiveness.	We have clarified that this report focused on clinical comparative effectiveness and that the cost of the various diagnostic strategies was not evaluated.
Peer Reviewer 6	General	The target population and audience are defined, but could potentially be more explicitly stated. A natural audience for this report could be researchers, grantmakers, regulators, and practicing clinicians among others. The authors describe the target population of this report as women with a chest pain syndrome. It was less clear if other symptomatology in women is also included or only "chest syndromes." I think the population that the review is attempting to capture are those women with symptoms for which a practicing clinician would like to evaluate CAD or ischemia as an etiology. The sentences on p. 2, beginning at line 23 describe the ACC/AHA recommendations about the population for diagnostic testing, but perhaps a clearer statement defining this as the target population for the review or defining the risk for CAD (e.g., intermediate risk) would be more explicit.	We have clarified that the target population was women with symptoms suspicious for CAD (or suspected CAD) and that this is an intermediate risk population.
Peer Reviewer 6	General	Clarity and Usability: The main points are quite clearly presented. Essentially, more research is needed for head to head comparisons of existing and new modalities for non-invasive testing in women with chest syndromes. I believe that these conclusions can be used as the basis to inform grant-making, research, and practice decisions.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Feinstein, Steven (International Contrast Ultrasound Society)	General	On behalf of the International Contrast Ultrasound Society ("ICUS"), we appreciate the opportunity to provide comments to the Agency for Healthcare Research and Quality on the research review draft entitled "Comparative Effectiveness of Noninvasive Technologies for the Diagnosis of Coronary Artery Disease in Women." ICUS is the world's only professional society exclusively devoted to the safe and appropriate use of contrast-enhanced ultrasound ("CEUS") medical imaging. ICUS represents physicians and scientists from the fields of cardiology, radiology, vascular medicine, gastro-intestinal medicine, oncology, and other medical specialties from North, Central and South America; Europe; Asia; Africa; and the Middle East.	Noted
Feinstein, Steven (International Contrast Ultrasound Society)	General	ICUS believes that no woman should be exposed to unnecessary ionizing radiation in connection with diagnostic imaging for the assessment of cardiovascular disease (or other medical conditions). According to the Joint Commission's latest Sentinel Alert, the risks associated with the use of ionizing radiation in diagnostic imaging include cancer, burns and other injuries, and over the past two decades the American public's total exposure to ionizing radiation has nearly doubled.1 The Joint Commission recommends that steps be taken "to eliminate avoidable exposure to radiation," with ultrasound or MRI used whenever these tests will produce the required diagnostic information at a similar quality level. Women are more biologically vulnerable to the effects of radiation than men, and this risk increases with decreasing age; many women who are referred for stress testing are in this "younger, physiologically vulnerable" demographic.	We agree that ionizing radiation is a risk for some diagnostic imaging modalities.
Feinstein, Steven (International Contrast Ultrasound Society)	General	CEUS is a form of ultrasound that uses an ultrasound contrast agent to improve the clarity and accuracy of a standard echocardiogram. A CEUS stress echocardiogram produces diagnostic information at a quality level similar to that of a SPECT study, but is performed without exposing patients to any ionizing radiation whatsoever and without the type of dye that can damage kidneys or other organs. The safety and diagnostic reliability of CEUS is well established in studies of large populations that included women, without any indication that the results are not applicable to women in particular. As previously communicated in our July 7, 2010, letter to AHRQ and in other settings, ICUS strongly believes that, at a time of heightened concern over unnecessary exposure to ionizing radiation as a result of medical imaging, an ionizing-radiation free tool such as CEUS should be considered as part of any comparative effectiveness review of noninvasive technologies for the diagnosis of coronary artery disease in women.	We included CEUS in the literature search and found no studies that reported results for women separately.

Commentator & Affiliation	Section	Comment	Response
Feinstein, Steven (International Contrast Ultrasound Society)	General	While the draft research review asserts that it includes studies of "good quality," it does not currently include studies of CEUS with stress echo. This omission occurs despite the authors' apparent acknowledgement of the possibility of using CEUS with stress echo, and that CEUS can improve the quality of standard echocardiograms. A 2009 study found that the appropriate use of CEUS improved endocardial visualization, which, in turn, positively affected diagnostic efficiency and resource utilization and improved patient management.2 Indeed, Page 50 of the draft research review reflects the omission of CEUS with stress echo, stating: Overall, within a given testing modality, the summary sensitivities and specificities were similar for both types of populations (known and no known CAD) and for all studies when compared with good-quality studies. When accounting for only the good-quality studies, it appeared that the diagnostic accuracy of detecting CAD in women was better (in descending order) for CTA, SPECT, ECHO, CMR, and ECG.	Reference 2 was not included in this review for the following reasons: (1) It did not include women with symptomatic chest pain, (2) all women were known to have CAD, (3) study had no comparison to another NIT or cardiac angiography, (4) data for women was not reported as a subgroup, (5) study did not have outcomes of interest for the key questions. Appendix E of the full report contains a list of all excluded studies with reasons for exclusion.
Feinstein, Steven (International Contrast Ultrasound Society)	General	The omission of CEUS with stress echo is particularly troublesome in light of the June 2010 Ontario Health Technology Assessment3, which, citing over 39 CEUS studies associated with stress echo testing, concluded that: 1. Stress ECHO with contrast has a higher diagnostic accuracy in the diagnosis of CAD than stress ECHO (without contrast). 2. Stress ECHO with contrast seems to have a similar diagnostic accuracy to 99 technetium SPECT. 3. The addition of contrast to ECHO in patients with previous suboptimal ECHO results significantly improves interpretability of the results. 4. Statistically, the addition of contrast agents to stress ECO tests does not significantly improve patient mortality rates.	We have reviewed the studies included in this report (Medical Advisory Secretariat, "Stress echocardiography with contrast for the diagnosis of coronary artery disease, an evidence-based analysis," Ontario Health Technology Assessment Series [Internet] 2010 June, Vol. 10, No. 10.) and did not find any articles that met our eligibility criteria.
Feinstein, Steven (International Contrast Ultrasound Society)	General	One specific CEUS stress echo study concluded that there was no short-term or long-term risk of death or myocardial infarction associated with CEUS among male and female patients. The safety during stress echo was shown with both genders (Males/females in contrast cohort = 57.4% / 42.6%, while males / female in non contrast cohort = 52.8% / 47.2%).4	This reference did not meet our eligibility criteria since it compared stress echo with and without contrast and did not include cardiac catheterization as the gold standard. (Abdelmoneim SS, Bernier M, Scott CG, et al. Safety of contrast agent use during stress echocardiography: a 4-year experience from a single-center cohort study of 26,774 patients. JACC Cardiovasc Imaging 2009; 2:1048-56. The report stated that 225,008 cardiac patients worldwide, carefully documented, have received echo contrast without evidence of increased death or serious cardiac event.)

Commentator & Affiliation	Section	Comment	Response
Feinstein, Steven (International Contrast Ultrasound Society)	General	Another specific study concluded that CEUS with stress echo should be the preferred diagnostic test when evaluating chest pain in women of low to intermediate probability of coronary disease. This study found that there appears to be no compromise in the sensitivity of CEUS with stress echo when compared to gated spect stress sestamibi in this population.5	The reference mentioned by the reviewer was only an abstract. Only peer-reviewed publications were included in this report. Ford K, McElroy B, Gwinn J, and Garneau R. A Comparison of Contrast Stress Echo and Stress Sestamibi in Women. Journal of the American Society of Echocardiography 2005:18:505. Abstract presented at the Sixteenth Annual Scientific Sessions, American Society of Echocardiography 2005, Boston, Massachusetts
Feinstein, Steven (International Contrast Ultrasound Society)	General	The Intersocietal Commission for the Accreditation of Echocardiography Laboratories ("ICAEL") is dedicated to ensuring quality patient care within the medical specialty of echocardiography. In keeping with this mission, the ICAEL develops and provides facility accreditation programs for echocardiography testing. A facility's ICAEL accreditation is contingent on whether it is in substantial compliance with the ICAEL Standards, which set forth the "minimal requirements for echocardiography laboratories to provide high quality care."6 Since December 2010, ICAEL has included contrast enhanced ultrasound in its Standards. In doing so, ICAEL has underscored the important role that ultrasound contrast agents play in improving accuracy. ICAEL recommends, and in some cases requires, the use of contrast agents in certain circumstances.	Noted
Feinstein, Steven (International Contrast Ultrasound Society)	General	Finally, in an effort to "improve patient care and health outcomes in a cost effective manner", The American College of Cardiology, the American Heart Association, the American Society of Echocardiography, and other prominent professional societies have published Appropriate Use Criteria for a variety of cardiovascular testing modalities. In 2011, a combined and updated document was published for transthoracic echocardiography, transesophageal echocardiography, and stress echocardiography. The AUC document indicates use of an ultrasound contrast agent is appropriate when ">=2 contiguous myocardial segments are not seen on non-contrast images."7	Noted

Commentator & Affiliation	Section	Comment	Response
Feinstein, Steven (International Contrast Ultrasound Society)	General	The inclusion of CEUS in AHRQ's ongoing comparative effectiveness study will improve the quality of the results of the study and avoid omitting a scientifically proven tool that could greatly enhance the safety of diagnostic procedures in women while avoiding exposure to ionizing radiation. Thank you in advance for your consideration of the foregoing. ICUS and its members would be pleased to answer any questions or provide any further information that you might find helpful.	Thank you for reviewing the report and bringing these references to our attention.
Feinstein, Steven (International Contrast Ultrasound Society)	General	1 Radiation Risks of Diagnostic Imaging, The Joint Commission Sentinel Event Alert, Issue 47, August 24, 2011. 2 Mustafa Kurt, MD, et al., Department of Cardiology, The Methodist Hospital – J Am Coll Cardiol, 2009; 53:802-810, expedited online publication 11 February 2009, © 2009 by the American College of Cardiology Foundation 3 Medical Advisory Secretariat, "Stress echocardiography with contrast for the diagnosis of coronary artery disease, an evidence-based analysis," Ontario Health Technology Assessment Series (Internet) 2010 June, Vol. 10, No. 10. 4 Abdelmoneim SS, Bernier M, Scott CG, et al. Safety of contrast agent use during stress echocardiography: a 4-year experience from a single-center cohort study of 26,774 patients. JACC Cardiovasc Imaging 2009; 2:1048-56. The report stated that 225,008 cardiac patients world-wide, carefully documented, have received echo contrast without evidence of increased death or serious cardiac event. 5 Ford K, McElroy B, Gwinn J, and Garneau R. A Comparison of Contrast Stress Echo and Stress Sestamibi in Women. Journal of the American Society of Echocardiography 2005:18:505. Abstract presented at the Sixteenth Annual Scientific Sessions, American Society of Echocardiography 2005, Boston, Massachusetts. 6 http://www.icael.org/icael/main/what_is_accreditation.htm 7 Douglas PS, Garcia MJ, Haines DE, Lai WW, Manning WJ, Patel AR, Picard MH, Polk DM, Ragosta M, Ward RP, Weiner RB. ACCF/ASE/AHA/ASNC/HFSA/HRS/SCAI/SCCM/SCCT/SCMR 2011 appropriate use criteria for echocardiography, Journal of the American College of Cardiology (2010), doi:10.1016/j.jacc.2010.11.002.	All references were reviewed for potential inclusion in the report and none met the eligibility criteria for the reasons stated in the responses above.

Commentator & Affiliation	Section	Comment	Response
Fisher, Dave (Medical Imaging and Technology Alliance)	General	The Medical Imaging and Technology Alliance (MITA) appreciates this opportunity to comment on the Agency for Healthcare Research and Quality's (AHRQ) draft research review entitled the "Comparative Effectiveness of Noninvasive Technologies for the Diagnosis of Coronary Artery Disease in Women."1 As the leading trade association representing medical imaging and radiotherapy technology manufacturers, we have an in-depth understanding of the significant benefits to the health of Americans that medical imaging, radiotherapy and proton therapy provides. MITA has long supported the efficient use of imaging equipment and looks forward to working with you to continue to improve Americans' healthcare through appropriate use of these technologies for the early detection, diagnosis, staging, therapy monitoring, and surveillance of many diseases.	Noted
Fisher, Dave (Medical Imaging and Technology Alliance)	General	Medical imaging procedures are usually non-invasive and encompasses X-ray imaging, computed tomography (CT) scans, radiation therapy, related image acquisitions, diagnostic ultrasound, and nuclear medical imaging (including positron emission tomography (PET)), and magnetic resonance imaging (MRI). Medical imaging is used to diagnose patients with disease, often reducing the need for costly medical services and invasive surgical procedures.2 In addition, medical imaging equipment often is used to select, guide and facilitate effective treatment, for example, by using image guidance for surgical or radiotherapeutic interventions.3	Noted
Fisher, Dave (Medical Imaging and Technology Alliance)	General	Specific to this draft study, imaging is used as a main tool for the diagnosis of women with suspected coronary artery disease through the use of ECHO, SPECT, CTA and CMR procedures. These procedures are central to the diagnosis of coronary artery disease in both men and women, and MITA appreciates your thorough and thoughtful consideration of these important tools.	Thank you.
Fisher, Dave (Medical Imaging and Technology Alliance)	General	MITA appreciates the effort that AHRQ invested to bring about this report and has found the overall content to be well prepared and thorough. We would however, like to stress a few areas for future consideration. First, we feel that the body of evidence on this technology has been well established. However, as you know, imaging technology benefits from continual advancement. For example, recent low-dose technology has altered the risk-benefit assessment for Coronary CT Angiography. Therefore, we ask that this study be updated to reflect advancements in patient care as new technologies and evidence are made available.	We recognize that technology is continually changing, with likely effects on risk-benefit assessment for any given test. This report was updated to incorporate information from recent publications; advancements in patient care as reported in eligible studies were reflected in the final report. We have added clarification that recent advancements in technology have reduced the radiation exposure for coronary CTA.

Commentator & Affiliation	Section	Comment	Response
Fisher, Dave (Medical Imaging and Technology Alliance)	General (ES-20, ES- 21)	Second, we ask that AHRQ support additional studies of the effectiveness of noninvasive technologies for the diagnosis of coronary artery disease related specifically to women. The report notes that 78 percent of 1,452 citations were excluded because they did not report data on women separately. The report observes, "it would aid future comparisons of modalities of study authors were to report the primary data for women and men separately." Further, as noted in the report, the predictive value of diagnostic studies in women with chest pain is different in men. We feel that additional research to understand this difference and value that specific technologies may bring to women merit additional study.	Noted
Fisher, Dave (Medical Imaging and Technology Alliance)	General	Last, we would ask that AHRQ also support the study and understanding of the relative values of the various modalities. As there are assigned values in the report, it would be beneficial to practitioners to have a greater depth of understanding of these values in order to best apply technologies to most appropriately treat their patients. Again we appreciate this opportunity to submit our comments. If you have any questions or would like to discuss these matters further, please contact me at 703-841-3279. Thank you for consideration of these comments.	Noted; thank you for your comments.
Fisher, Dave (Medical Imaging and Technology Alliance)	General	1 AHRQ, Comparative Effectiveness of Noninvasive Technologies for the Diagnosis of Coronary Artery Disease in Women, Draft, August 30,, 2011, http://www.effectivehealthcare.ahrq.gov/ehc/products/202/770/NIT_Draft-Report_20110829.pdf (hereinafter "Draft Report"). 2 See, e.g., Multidetector-Row Computed Tomography in Suspected Pulmonary Embolism," Perrier, et. al., New England Journal of Medicine, Vol 352, No 17; pp1760-1768, April 28, 2005. 3 See, e.g., Jelinek, JS et al. "Diagnosis of Primary Bone Tumors with Image-Guided Percutaneous Biopsy: Experience with 110 Tumors." Radiology. 223 (2002): 731 - 737.	References 2 and 3 were reviewed for findings potentially relevant for this report and did not meet our eligibility criteria.
Flamm, Scott (Society for Cardiovascular Magnetic Resonance)	General	The Society for Cardiovascular Magnetic Resonance (SCMR) is pleased to offer comments on the Agency for Healthcare Research & Quality's (AHRQ) draft report on the Comparative Effectiveness of Noninvasive Technologies for the Diagnosis of Coronary Artery Disease in Women. SCMR is the recognized representative and advocate for physicians, scientists, and technologists who work in the field of cardiovascular magnetic resonance (CMR).	Noted

Commentator & Affiliation	Section	Comment	Response
Flamm, Scott (Society for Cardiovascular Magnetic Resonance)	General	SCMR is very concerned about the quality, accuracy, and inclusiveness of the data reviews and interpretation of data throughout this report. We have identified numerous weaknesses and errors in the review of literature cited in the draft report, and we have noted below more contemporary references for CMR that have been omitted from the draft. Although our specific comments will address only CMR, we believe that there are also serious flaws and weaknesses in the reviews, identification of references, and interpretation of data for other cardiovascular imaging modalities. We encourage AHRQ to carefully consider the comments of our sister societies as well, in order to more accurately understand and evaluate the capabilities of various cardiovascular imaging modalities.	Noted
Flamm, Scott (Society for Cardiovascular Magnetic Resonance)	General	1) The literature review included in this report is not up to date, and it appears that rather than a full review of the literature, the studies included in this report were highly selective and skewed toward old studies of CMR. Given the rapid technical development in the field in the past decade, such older studies are not representative of CMR capabilities at the present time. This review included five relatively small CMR papers (totaling ~ 500 patients only) from mixed populations dating to 2003 when CMR was a very different technology and did not include the most relevant papers that specifically addressed diagnosis and prognosis of women. There are now 17 prognostic studies that involve stress CMR totaling more than 7000 patients for diagnosis and prognosis of CAD.	We have reviewed the references and give a detailed response below.

Commentator & Affiliation	Section	Comment	Response
Flamm, Scott (Society for Cardiovascular Magnetic Resonance)	General	Examples of recent CMR papers that targeted diagnosis and prognosis of women include Wallace E et al (N=266), JACC Cardiovasc Imaging. 2009 Mar;2(3):299-307 and Coelho-Filho O et al (N=424) in JACC Cardiovasc Imaging 2011 Aug;4(8):850-61. In addition, there are two excellent meta-analyses (Nandalur et al JACC 2007;50:1343-53 [Imaging of stress-induced wall motion abnormalities demonstrated a sensitivity of 0.83 (95% confidence interval [CI] 0.79 to 0.88) and specificity of 0.86 (95% CI 0.81 to 0.91) on a patient level (disease prevalence = 70.5%). Perfusion imaging demonstrated a sensitivity of 0.91 (95% CI 0.88 to 0.94) and specificity of 0.81 (95% CI 0.77 to 0.85) on a patient level (disease prevalence = 57.4%] and Hamon et al JCMR 2010;12:29 [The overall patient-based analysis demonstrated a sensitivity of 89% (95% CI: 88-91%), and a specificity of 80% (95% CI: 78-83%). Adenosine stress perfusion CMR had better sensitivity than dipyridamole stress CMR(90% (88-92%) versus 86% (80-90%), P = 0.022), and a tendency to a better specificity (81% (78-84%) versus 77% (71-82%), P = 0.065).]) that were included only in the appendix, with results that do not appear to have been considered.	Thank you for discussing these articles. The Wallace 2009 reference did not meet our criteria for inclusion since it did not compare one NIT to another or to diagnostic coronary angiography. In addition, the results of the symptomatic subgroup were not reported separately. The Coelho-Filho 2011 reference has been included in the final report. While systematic reviews and meta-analyses did not meet our criteria to be directly abstracted, we did review the component references of such articles to identify additional articles reporting original trial data that could be relevant to this review. Such articles were then screened to determine if they met criteria for inclusion. Component references of both the Nandalur 2007 and Hamon 2010 meta-analyses were reviewed in this fashion.
Flamm, Scott (Society for Cardiovascular Magnetic Resonance)	General	2) There appears to be a serious discrepancy in the review and analysis of at least one of the studies selected for inclusion in the draft report. In the analysis AHRQ shows an SROC that quotes Klem's paper in JACC Cardiovasc Imaging2008 Jul;1(4):436-45, giving a sensitivity of 70%, but the paper itself reports a sensitivity of 84%. We do not understand this discrepancy. See abstract below:	This study reports its findings with % stenosis defined to be positive for CAD being either >70% and >50%. The latter was what we used in our analyses as a definition—and under this definition, the sensitivity was 70%.

Commentator & Affiliation	Section	Comment	Response
Flamm, Scott (Society for Cardiovascular Magnetic Resonance)	General	Abstract OBJECTIVES: We wanted to assess the value of cardiovascular magnetic resonance (CMR) stress testing for evaluation of women with suspected coronary artery disease (CAD). BACKGROUND: A combined perfusion and infarction CMR examination can accurately diagnose CAD in the clinical setting in a mixed gender population. METHODS: We prospectively enrolled 147 consecutive women with chest pain or other symptoms suggestive of CAD at 2 centers (Duke University Medical Center, Robert-Bosch-Krankenhaus). Each patient underwent a comprehensive clinical evaluation, a CMR stress test consisting of cine rest function, adenosine-stress and rest perfusion, and delayed-enhancement CMR infarction imaging, and X-ray coronary angiography within 24 h. The components of the CMR test were analyzed visually both in isolation and combined using a pre-specified algorithm. Coronary artery disease was defined as stenosis > or =70% on quantitative analysis of coronary angiography. RESULTS: Cardiovascular magnetic resonance imaging was completed in 136 females (63.0 +/- 11.1 years), 37 (27%) women had CAD on coronary angiography. The combined CMR stress test had a sensitivity, specificity, and accuracy of 84%, 88%, and 87%, respectively, for the diagnosis of CAD. Diagnostic accuracy was high at both sites (Duke University Medical Center 82%, Robert-Bosch-Krankenhaus 90%; p = 0.18). The accuracy for the detection of CAD was reduced when intermediate grade stenoses were included (82% vs. 87%; p = 0.01 compared the cutoff of stenosis > or =50% vs. > or =70%). The sensitivity was lower in women with single-vessel disease (71% vs. 100%; p = 0.06 compared with multivessel disease) and small left ventricular mass (69% vs. 95%; p = 0.04 for left ventricular mass < or =97 g vs. >97 g). The latter difference was even more significant after accounting for end-diastolic volumes (70% vs. 100%; p = 0.02 for left ventricular mass indexed to end-diastolic volume < or =1.15 g/ml vs. >1.15 g/ml). CONCLUSIONS: A multicomponent CMR stress test can accurately diagnose CAD in women. Detection of CAD in women with intermediate grade stenosis, single-vessel disease, and with small hearts is challenging.	Please refer to our response above.

Commentator & Affiliation	Section	Comment	Response
Flamm, Scott (Society for Cardiovascular Magnetic Resonance)	General	SCMR urges AHRQ to carefully evaluate the most recent studies of CMR. If these data are not accurately reviewed and reported, other scientific reviews will also be suspect and called into question even if accurate. We also believe that an up to date CMR evaluation would provide a clear demonstration of its appropriateness for clinical use in women.	Per AHRQ methodology, we updated the literature search during the public posting of this document up to September 12, 2011. We reviewed approximately 265 articles of CMR, and of these, 7 met our inclusion criteria. We encourage the reporting of results in the women subgroup for all trials of CMR.
Holmes, David (American College of Cardiology)	General	The American College of Cardiology (ACC) is pleased to submit comments on the draft report regarding Comparative Effectiveness of Noninvasive Technologies for the Diagnosis of Coronary Artery Disease in Women. The College is a 39,000-member nonprofit medical society composed of physicians, nurses, nurse practitioners, physician assistants, pharmacists, and practice managers, and bestows credentials upon cardiovascular specialists who meet its stringent qualifications. The ACC is a leader in the formulation of health policy, standards and guidelines, and is a staunch supporter of cardiovascular research. We would like to thank the Effective Health Care Program for its work on this report.	Noted
Holmes, David (American College of Cardiology)	General	We understand several other medical societies with expertise in cardiovascular imaging have commented with concerns about the draft. We urge you to carefully examine those very thorough comments. We are also providing some additional suggestions that we believe could improve the report.	Thank you.
Holmes, David (American College of Cardiology)	General	We found the report to focus on only a select portion of the large amount of literature available on the effectiveness of noninvasive technologies (NITs) for assessment of coronary artery disease in women. To be included in this document, one NIT had to be compared to another NIT or to coronary angiography. It is well recognized in clinical practice that not all patients who undergo stress tests also undergo coronary angiography. We suggest review of the following additional references:	Studies were required to have a comparative diagnostic modality for this comparative effectiveness report. We agree that, in routine clinical practice, only a subset of patients undergo coronary angiography.

Commentator & Affiliation	Section	Comment	Response
Holmes, David (American College of Cardiology)	General	Ghostine S, Caussin C, Daoud B, et al. Noninvasive detection of coronary artery disease in patients with left bundle branch block using 64-slice computed tomography. J Am Coll Cardiol. 2006;48:1929:1934. Motoyama S, Anno H, Sarai M, et al. Noninvasive coronary angiography with a prototype 256-row area detector computed tomography system: comparison with conventional invasive coronary angiography. J Am Coll Cardiol. 2008;51:773–5. Budoff MJ, Dowe D, Jollis JG, et al. Diagnostic performance of 64-multidetector row coronary computed tomographic angiography for evaluation of coronary artery stenosis in individuals without known coronary artery disease: results from the prospective multicenter ACCURACY (Assessment by Coronary Computed Tomographic Angiography of Individuals Undergoing Invasive Coronary Angiography) trial. J Am Coll Cardiol. 2008;52:1724 –32. Meijboom WB, Meijs MF, Schuijf JD, et al. Diagnostic accuracy of 64-slice computed tomography coronary angiography: a prospective, multicenter, multivendor study. J Am Coll Cardiol. 2008;52:2135–44. Cademartiri et al. Diagnostic accuracy of 64-slice computed tomography coronary angiography in patients with low-to-intermediate risk. Radiol Med. 2007;112:969–81. Zhang et al. Diagnostic accuracy of dual-source CT coronary angiography: The effect of average heart rate, heart rate variability, and calcium score in a clinical perspective. Act Radiologica. 2010;7:727-40. Nasis et al. Diagnostic Accuracy of Noninvasive Coronary Angiography With 320- Detector Row Computed Tomography Am J Cardiol. 2010;106:1429 –35.	Ghostine 2006 was excluded for not including women-specific data. Motoyama 2008 was excluded for not being a clinical study report. Budoff 2008 was excluded for not including women-specific data. Meijboom 2008 is a paper related to Meijboom 2007 (Meijboom WB, Weustink AC, Pugliese F, et al. Comparison of diagnostic accuracy of 64-slice computed tomography coronary angiography in women versus men with angina pectoris. Am J Cardiol 2007;100(10):1532-7. PMID: 17996514) and Dharampal 2011 (Dharampal AS, Rossi A, Papadopoulou SL, et al. Is there a difference in the diagnostic accuracy of computed tomography coronary angiography between women and men? Coronary Artery Disease 2011;22(6):421-427. PMID: 2011455721), which both report results for the same prospective cohort, so we have included Dharampal 2011 in this report and have listed Meijboom 2007 and Meijboom 2008 as related papers. Cademartiri 2007 was excluded for not including women-specific data. Zhang 2010 and Nasis 2010 did not include women-specific data.
Holmes, David (American College of Cardiology)	General	Additionally, there are hundreds of articles on coronary computed tomographic angiography (CCTA) diagnostic accuracy, many of them available in the 2010 expert consensus document on CCTA. (Mark DB et al. ACCF/ACR/AHA/NASCI/SAIP/SCAI/ SCCT 2010 Expert Consensus Document on Coronary Computerized Tomographic Angiography: A Report of the American College of Cardiology Foundation Task Force on Expert Consensus Documents. Circulation. 2010;121:2509-43) The added power of more studies should produce more cogent conclusions regarding the utility of CCTA in women.	We have reviewed the references, and none met eligibility criteria.

Commentator & Affiliation	Section	Comment	Response
Holmes, David (American College of Cardiology)	General	The conclusion that "There were insufficient data to demonstrate that the use of specific NITs (compared with coronary angiography) routinely provided incremental risk stratification, prognostic information, or other meaningful information to improve decisionmaking," is concerning. Does this mean that these tests should not be performed in the clinical environment? Do the authors recommend that all patients with suspected coronary artery disease should undergo coronary angiography? It is unfortunate that comparative studies examining the various NIT testing modalities in women in terms of outcomes, treatment decisions, and harms do not exist. Part of the issue is that newer technologies such as coronary tomography angiography and coronary magnetic resonance imaging are not as widespread in use, and have not been available until 3 recently. We hope the draft's conclusion can be used as impetus to increase government funding to address these important issues.	We agree that the current state of the evidence base is lacking and should be used as an impetus to spur innovation and further research. Additionally, we acknowledge that our exclusion of noncomparator trials does not correspond with the clinical recommendations that are often provided by professional societies. The purpose of this review was to describe what was reported in the "comparative" literature—not to make recommendations about whether the tests should be performed or if coronary angiography should be done in all cases.
Min, James (Society of Cardiovascular Computed Tomography)	General	The Society of Cardiovascular Computed Tomography (SCCT) is pleased to offer comments on the Agency for Healthcare Research & Quality's (AHRQ) draft report on the Comparative Effectiveness of Noninvasive Technologies for the Diagnosis of Coronary Artery Disease in Women. SCCT, with nearly 3,000 members, is the professional society representing physicians, scientists and technologists advocating for research, education, and clinical excellence in the use of cardiovascular computed tomography.	Noted
Min, James (Society of Cardiovascular Computed Tomography)	General	In general, SCCT is concerned about the quality, accuracy, and inclusiveness of the data reviews and interpretation of data throughout this report. We have identified weaknesses in the review of literature cited in the draft report, and we have noted that numerous references for cardiac CT have been omitted from the draft. We encourage AHRQ to carefully consider the comments of our sister societies as well, in order to more accurately understand and evaluate the capabilities of various cardiovascular imaging techniques.	Noted; see the following response to the literature.
Min, James (Society of Cardiovascular Computed Tomography)	General	Our specific comments regarding the review of cardiac computed tomography are as follows: 1) Missing References Many key references related to coronary CT angiography are missing from this draft report, including the multicenter trials performed (ie – Accuracy Study, Meijboom, etc). Studies that should be reviewed, analyzed and included in the AHRQ report are:	Ghostine 2006 was excluded for not including women-specific data. Motoyama 2008 was excluded for not being a clinical study report. Budoff 2008 was excluded for not including women-specific data. Meijboom 2008 is a paper related to Meijboom

Commentator & Affiliation	Section	Comment	Response
		Ghostine S, Caussin C, Daoud B, et al. Noninvasive detection of coronary artery disease in patients with left bundle branch block using 64-slice computed tomography. J Am Coll Cardiol. 2006;48: 1929–34. Motoyama S, Anno H, Sarai M, et al. Noninvasive coronary angiography with a prototype 256-row area detector computed tomography system: comparison with conventional invasive coronary angiography. J Am Coll Cardiol. 2008;51:773–5. Budoff MJ, Dowe D, Jollis JG, et al. Diagnostic performance of 64-multidetector row coronary computed tomographic angiography for evaluation of coronary artery stenosis in individuals without known coronary artery disease: results from the prospective multicenter ACCURACY (Assessment by Coronary Computed Tomographic Angiography of Individuals Undergoing Invasive Coronary Angiography) trial. J Am Coll Cardiol. 2008;52:1724 –32. Meijboom WB, Meijs MF, Schuijf JD, et al. Diagnostic accuracy of 64-slice computed tomography coronary angiography: a prospective, multicenter, multivendor study. J Am Coll Cardiol. 2008;52:2135–44. Cademartiri et al. Diagnostic accuracy of 64-slice computed tomography coronary angiography in patients with low-to-intermediate risk Radiol med (2007) 112:969–981. Zhang et al. Diagnostic accuracy of dual-source CT coronary angiography: The effect of average heart rate, heart rate variability, and calcium score in a clinical perspective. ACTA RADIOLOGICA 2010:7:727-740. Nasis et al. Diagnostic Accuracy of Noninvasive Coronary Angiography With 320-Detector Row Computed Tomography Am J Cardiol 2010;106:1429 –1435.	2007 (Meijboom WB, Weustink AC, Pugliese F, et al. Comparison of diagnostic accuracy of 64-slice computed tomography coronary angiography in women versus men with angina pectoris. Am J Cardiol 2007;100(10):1532-7. PMID: 17996514) and Dharampal 2011 (Dharampal AS, Rossi A, Papadopoulou SL, et al. Is there a difference in the diagnostic accuracy of computed tomography coronary angiography between women and men? Coronary Artery Disease 2011;22(6):421-427. PMID: 2011455721.), which both report results for the same prospective cohort, so we have included Meijboom 2008 and Dharampal 2011 in this report. Cademartiri 2007 was excluded for not including women-specific data. Zhang 2010 and Nasis 2010 did not include women-specific data.
Min, James (Society of Cardiovascular Computed Tomography)	General	There are hundreds of articles on the diagnostic accuracy of coronary CT angriography , many of them available on the Mark DB et al Expert Consensus Document on CTA from 2010. The added power of additional studies would lead to more cogent conclusions regarding the utility of coronary CT angiography in women.	We have reviewed the references, and none met eligibility criteria.

Commentator & Affiliation	Section	Comment	Response
Min, James (Society of Cardiovascular Computed Tomography)	General	2) The summary and discussion final paragraph related to women should include coronary CT angiography, as it only discusses the other four modalities. We find no reason or basis to exclude coronary CT angiography.	We have looked at the final paragraphs of the Discussion section and have already included CTA in this section.
Min, James (Society of Cardiovascular Computed Tomography)	General	3) We urge AHRQ to use the new standardized nomenclature for cardiac CT and coronary CT angiography, as detailed in the SCCT guideline: Standardized medical terminology for cardiac computed tomography: A report of the Society of Cardiovascular Computed Tomography (Journal of Cardiovascular Computed Tomography, 2011). http://www.scct.org/documents/JCCT427.pdf	Thank you for bringing the new nomenclature to our attention. We have reviewed this reference and updated the terminology in the report to state “coronary CTA.”
Pellikka, Patricia (American Society of Echocardiography)	General	This study focused on only a select portion of the large amount of literature available on the effectiveness of noninvasive technologies for the assessment of coronary artery disease in women. To be included in this document, a noninvasive technology had to be compared to another noninvasive technology or to coronary angiography. It is well recognized in clinical practice that not all patients having stress tests also undergo coronary angiography. In fact, generally only about 10% of patients undergoing noninvasive stress testing also undergo coronary angiography. Therefore, focusing only on this small subset of the population will lead to confusion about the true sensitivity and specificity of the test (Roger V: Circulation 1997). Similarly, requiring another comparative test also eliminates many studies that would have provided important insights regarding the utility of noninvasive technologies. Patients undergoing a clinically indicated exercise imaging test often have a non diagnostic exercise electrocardiogram and therefore, comparing the ECG and the imaging modality will poorly represent the value of the exercise ECG. Studies comparing multiple imaging modalities generally include limited numbers of patients as such studies are costly and may involve significant radiation burden to the patient.	The focus of the review was on comparative effectiveness; therefore, the body of literature required a comparative diagnostic modality. We acknowledge that routine clinical care does not require two NIT modalities or an NIT modality and coronary angiography.
Pellikka, Patricia (American Society of Echocardiography)	General	Some of the biases of post test referral to coronary angiography can be overcome by evaluating outcomes in all patients undergoing a particular noninvasive test. These data were not included in this report. I would recommend that the search criteria be modified to include the numerous large outcome studies of NIT modalities.	We have stated in the Limitation section of this report that noncomparator studies were not reviewed.

Commentator & Affiliation	Section	Comment	Response
Pellikka, Patricia (American Society of Echocardiography)	General	The conclusion "there were insufficient data to demonstrate that the use of specific NITs routinely provided incremental risk stratification, prognostic information, or other meaningful information to improve decision making" is very concerning. Does this mean that these tests should not be performed in the clinical environment, or if performed, not reimbursed? Do the authors recommend that all patients with suspected coronary artery disease should undergo coronary angiography, not to mention its risks. Noninvasive tests may be equally predictive or even better predictive of events during follow-up. (From AM, McCully RB, JASE 23:207-14, 2010) and (Sicari R, JACC 41:589-95, 2003).	We have stated the fact that "comparator" studies have insufficient data. We have stated in the Limitation section that noncomparator studies were not reviewed.
Shaw, Leslee (American Society of Nuclear Cardiology)	General	The American Society of Nuclear Cardiology (ASNC) is pleased to provide comments on the Agency for Healthcare Research & Quality's (AHRQ) draft report on the Comparative Effectiveness of Noninvasive Technologies for the Diagnosis of Coronary Artery Disease in Women.	Noted
Shaw, Leslee (American Society of Nuclear Cardiology)	General	ASNC is a 4,700 member professional medical society, which provides a variety of continuing medical education programs related to nuclear cardiology and cardiovascular computed tomography develops standards and guidelines for training and practice, promotes accreditation and certification within the nuclear cardiology field, and is a major advocate for furthering research and excellence in nuclear cardiology and cardiovascular computed tomography.	Noted
Shaw, Leslee (American Society of Nuclear Cardiology)	General	Based upon a review of the AHRQ's synthesis of evidence on diagnostic testing in women; notably gender differences in the diagnosis of coronary artery disease (CAD), we have identified significant weaknesses in the quality of the reviews and interpretation of the available data. As such we are providing comment in hopes of highlighting the following challenges with the document in its current version. We have outlined below our initial, more general comments on the document's assumptions followed by more specific, detailed comments.	Noted

Commentator & Affiliation	Section	Comment	Response
Shaw, Leslee (American Society of Nuclear Cardiology)	General	Our general comments are as follows: Throughout the document, the authors note the importance of coronary angiography as a core test for diagnosis. In the de novo evaluation of stable chest pain, very few women would qualify for direct angiography and this should be supported in this document. Many references are made in this document that the discerning of candidates for diagnostic testing is unknown. This is far from true where a pretest likelihood of CAD remains the cornerstone of this referral process. Although the authors refer to pretest risk, it would be important to highlight this as central to appropriate referral patterns. Accordingly, this document should refer to the American College of Cardiology Foundation's (ACCF) appropriate use taskforce for relevant diagnostic testing documents. The authors should also review the recent ACCF/AHA guideline on detection of high risk asymptomatic individuals (Greenland JACC 2010); as several notations in this report discuss testing in asymptomatic individuals.	We agree with the reviewers that an essential first step in the process of test referral is determining the pretest probability of disease for a given patient. It should be noted that this review focused on symptomatic patients, and as such the review of high-risk asymptomatic individuals in Greenland JACC 2010 is informative but not directly relevant to this analysis. Instead, we have cited the AHA/ACC stable angina guidelines and Appropriate Use Criteria documents for the various NIT modalities in our discussion of the pretest risk.
Shaw, Leslee (American Society of Nuclear Cardiology)	General	Further, the authors should take care to employ standardized terminology when referring to each of the reviewed modalities. For example, the authors write: Cardiac perfusion and stress magnetic resonance imaging (CMR). In fact, the cardiac perfusion is also part of the stress examination with MR. So, a more correct phrasing would be stress myocardial perfusion and wall motion MR. The correct terminology for computed tomography (CT) would be coronary CT angiography; as per the Society of Cardiovascular Computed Tomography. These titles should be applied consistently throughout the document. As part of this, there are numerous places where it appears a lack of expertise may lead to a "mixing of evidence." For example, the authors combine the rest ECG with an exercise ECG for documenting ischemia. We urge the authors to take care to note that the utility of the rest ECG in a stable population is largely to note resting ST-T wave abnormalities which preclude accurate interpretation of exertional changes. This is fundamental to the practice of stress ECG interpretation and should be included in this report. Another fundamental piece of information for discerning the use of an exercise versus pharmacologic stress imaging is the ability of a woman to perform maximal levels of exercise. Given the greater magnitude of functional disability in females, this should be discussed in the report.	We have replaced "cardiac perfusion and stress magnetic resonance imaging" with "stress myocardial perfusion and wall motion magnetic resonance imaging" as well as "computed tomography (CT)" with "coronary CT angiography (coronary CTA)." Resting ECG refers to multifunction cardiogram (Premier Heart), not to the routine ECG. This manufacturer had requested consideration as an NIT diagnostic modality when the key questions were posted on the AHRQ Web site. We have discussed the influence of functional capacity on NIT testing in the Introduction.

Commentator & Affiliation	Section	Comment	Response
Shaw, Leslee (American Society of Nuclear Cardiology)	General	ASNC is also concerned that the identification of prognostic series in the literature is of poor quality. There are numerous gender comparisons of prognosis in very large registries often including more than 1,000 women. The highlighted series are but a fraction of these reports and a more thorough review of the risk assessment data should be provided.	We agree that there are registry studies with prognostic information in the literature. Since registries are noncomparative studies, these were excluded from the review and have been identified as a limitation of the report.
Shaw, Leslee (American Society of Nuclear Cardiology)	General	In addition, the authors have chosen to include radiation safety as part of this report. While we appreciate this inclusion and applaud the effort, it should be noted for all relevant sections that cancer risk should be reported as projected and not observed. Moreover, the authors should report the high degree of uncertainty on the projected cancer risk. Specifically, the authors wrote that radiation safety concerns are different between women and men. This statement should be clarified and the authors should acknowledge that there is evidence that reproductive organs are more radiosensitive. As such, exposure to these tissues will result in a higher projected cancer risk. Moreover, this statement should also comment on not using tests with ionizing radiation in young women; unless deemed clinically necessary.	We have added more details about the radiation risk to the Safety Concerns and Risks section.
Shaw, Leslee (American Society of Nuclear Cardiology)	General	We also urge the authors to proceed with caution where they discuss diagnostic performance based on single, small series of women. For example, the authors write: From a limited number of studies, it appeared that: (1) the presence of LVH reduced the specificity of stress ECG, SPECT, and ECHO,(Page 10) (2) the use of beta blocker agents reduced the specificity of stress ECG and the sensitivity and specificity of SPECT,(Page 11); and (3) the positive predictive value increased as the pretest probability rose for ECG and ECHO. (Page 12). In the LVH example, this series includes 74 hypertensive women and an even smaller patient series with LVH. It is unclear from this report whether this is a group of symptomatic women; if not then this reference to this report should be deleted. Moreover, the example of the reduced accuracy with beta blockers is in 51 patients. Guidelines from the ASNC recommend withholding beta blockers prior to myocardial perfusion imaging and this would make the emphasis on reduced accuracy an invalid concern. Moreover, common sense would dictate that the use of beta blockers would affect all stress testing modalities.	We acknowledge that these statements are the result of small studies and have modified the summary to state that there is insufficient evidence on modifiers of diagnostic accuracy in women.

Commentator & Affiliation	Section	Comment	Response
Shaw, Leslee (American Society of Nuclear Cardiology)	General	We also note that in the summary of findings in Table ES-3, it appears that there is overlap and commonalities in the diagnostic accuracy of the imaging modalities. It is hard to imagine that small differences in sensitivity and specificity are meaningful. A clear message is the added accuracy of stress imaging tools over and above the ECG but that would be about all that can be reliably stated. Just because a statistical test may be employed, it does not mean that the results are meaningful.	We agree with the reviewer that caution should be used in drawing clinical conclusions from findings that may be statistically significant but may not be clinically meaningful. Therefore, we provided the overall point estimates for each study with the confidence intervals so that readers may view the data. We also provided information about how the specific accuracy of the tests translates into positive and negative predictive values given the underlying prevalence of disease. Additionally, we encourage the clinical community to perform additional outcomes studies so that these differences and their impact on important patient outcomes may be understood.

Commentator & Affiliation	Section	Comment	Response
Shaw, Leslee (American Society of Nuclear Cardiology)	General	In the nuclear imaging section, the authors note a high frequency of breast artifact in women. The frequency of breast artifact has been reduced substantially over the last 15 years due to the more common use of Technetium (Tc-99m) agents, attenuation correction algorithms, and supine imaging. This section should reflect contemporary nuclear cardiology practices that are commonly employed in millions of women annually. This section should also note the challenges in imaging women with large breasts or in obese patients for all imaging modalities. Moreover, the authors should directly site the literature on a reduced accuracy with perfusion imaging in women with single vessel CAD as this is unknown to our reviewers. However, the reduced accuracy with echocardiography has been reported in the American Heart Association (AHA) statement on diagnostic testing in women (Mieres Circulation 2005). Given the ischemic cascade and the sensitivity of a perfusion abnormality for detection of an intermediate stenosis, this statement on reduced accuracy does not make sense for myocardial perfusion imaging but more so for a wall motion abnormality. The authors should consider a discussion on the ischemic cascade and how a given abnormality (global ventricular function, wall motion abnormality, abnormal perfusion) would be detected in a temporal manner and what underlying disease would be more often detected. For example, a wall motion abnormality when noted is more often associated with a significant stenosis. This discussion would help readers to understand the pathophysiologic basis for coronary disease detection with common imaging markers of ischemia.	We have clarified that advances in nuclear imaging have reduced breast artifact. We acknowledge the rest of your comments.
Shaw, Leslee (American Society of Nuclear Cardiology)	General	Finally, we would like to point out to that coronary CT angiography is not an “emerging modality” as there are an estimated several million procedures performed annually. The AHA statement on CT is now 5 years old and not recent. Further, the comparison of CT with MR is odd; given that these modalities are quite different. As such, we believe it would be preferable to have two separate sections to address each of these modalities in isolation.	We have changed “emerging” to “anatomic.” We acknowledge that use of CTA and CMR are increasing over time, but the proportion of these remains small in comparison with the use of exercise ECG, ECHO, and SPECT; therefore, we have kept them in the same section of the Introduction.

Commentator & Affiliation	Section	Comment	Response
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	The conclusions of the report are relatively limited: The primary conclusion is that ECG stress testing is the only modality that has a statistically significant difference in women in terms of being less sensitive and less specific than the imaging modalities. The review includes 11 studies reporting safety concerns and risks, including three studies that show increased rates of cancer in women attributed to CTA. While the risk of cancer has been studies in association, it must also be noted that nuclear imaging frequently involves exposure to an equal if not greater amount of radiation than CTA. The authors stated on page 86 "Other than higher mean effective radiation doses for CTA studies for women compared with men, the extant literature does not provide sufficient evidence to conclude that safety concerns, risks, or radiation exposure associated with different NITs to diagnose CAD in patients with chest pain syndrome differ significantly between women and men." Given the know radiation exposure associated with nuclear imaging, the lack of data does not translate into the lack of risk. This cumulative exposure risk was reported by Eisenberg et al in CMAJ. (CMAJ February 7, 2011) In addition, the review includes 11 studies that assess risk stratification, clinical outcomes, and treatment decision-making, including 2 studies that demonstrated that an abnormal stress ECG or abnormal stress CMR, in the setting of normal invasive coronary angiography, resulted in a higher likelihood of future cardiac events.	Thank you for these comments. We have tried to use language in the report that differentiates between lack of data and lack of risk. The report by Eisenberg was not eligible for inclusion because it focused on patients with known CAD.
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	While we applaud AHRQ for undertaking a review of the clinical literature in this field, and while we agree that additional study of the comparative effectiveness of various NITs for the diagnosis of CAD in women is needed, we have a number of concerns about the Draft Report. First, and perhaps most importantly, the Draft Report focuses on only a select portion of the large volume of literature available on the effectiveness of noninvasive technologies for assessment of coronary artery disease in women. To be included in this document, a noninvasive technology had to be compared to another noninvasive technology or to coronary angiography. Of the studies included in the report it appears that all or virtually all compare a NIT to coronary angiography.	The report required direct comparisons of NIT modalities to each other or to coronary angiography. We did not include the noncomparator literature since it would require indirect comparisons of the NIT modalities.

Commentator & Affiliation	Section	Comment	Response
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	The selection criteria used in this analysis have a number of significant implications in terms of its findings regarding the comparative sensitivity and specificity of various NITs and in terms of their relative clinical utility. It is well recognized in clinical practice that not all patients having stress tests undergo coronary angiography. In fact, generally only about 10% of patients undergoing noninvasive stress testing also undergo coronary angiography. Therefore, focusing only on this small subset of the population will lead to confusion about the true sensitivity and specificity of the test (Roger V: Circulation 1997). Some of the biases inherent in focusing on this small subset of the population can be overcome by evaluating outcomes in all patients undergoing a particular NIT; however, these data were not included in the report.	The focus of the review was on comparative effectiveness; therefore, the body of literature required a comparative diagnostic modality. We acknowledge that routine clinical care does not require two NIT modalities or an NIT modality and coronary angiography.
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	Moreover, the selection criteria also appear to have eliminated many studies that would have provided important insights regarding the relative clinical utility of NITs. Patients undergoing a clinically indicated exercise imaging test often have a non-diagnostic exercise electrocardiogram and therefore, comparing the ECG and the imaging modality will poorly represent the value of the exercise ECG. Studies comparing multiple imaging modalities generally include limited numbers of patients, as such studies are costly and may involve significant radiation burden to the patient, and so may not have met the authors' rigorous selection criteria. However, such studies can provide important clinical insights.	The focus of the review was on comparative effectiveness; therefore, the body of literature required a comparative diagnostic modality. We acknowledge that routine clinical care does not require two NIT modalities or an NIT modality and coronary angiography.
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	To be certain, it appears that echocardiography has a number of advantages over NITs for the diagnosis of CAD in women, as summarized in the attached President's Communication form JASE, the Journal of the American Society of Echocardiography. We also recommend an excellent review article on stress echocardiography by Gottdiener (Gottdiener JS. Progress in Cardiovascular Diseases. 2001;43:315-334). For these reasons, we would strongly recommend that the search criteria be modified to include the numerous large outcome studies of NIT modalities.	The search criteria will remain the same; i.e., require direct comparison of NIT modalities with each other or with coronary angiography.

Commentator & Affiliation	Section	Comment	Response
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	In fact, it is unclear to us whether an analysis of studies comparing NITs to coronary angiography is appropriate to answer on the most important clinical questions posed in the Draft Report: Is there evidence that the use of NITs (when compared with other NITs or with coronary angiography) in women improves: KQ 3a. Risk stratification/prognostic information? KQ 3b. Decisionmaking regarding treatment options (e.g., revascularization, optimal medical therapy)? KQ 3c. Clinical outcomes (e.g., death, myocardial infarction, unstable angina, hospitalization, revascularization, angina relief, quality of life)?	We have acknowledged in the Limitation section that noncomparator studies often contain information that can address these questions.
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	Women, in particular, are known to have higher incidences of chest pain in the setting of normal coronary arteries, which in many situations, the of the chest pain, is thought to be due to microvascular coronary disease. The limitations of coronary angiography in visualizing the microvasculature is well described in the literature as are the risks of the procedure. The comparison of echocardiography to coronary angiography in women is fraught with difficulty given the recently identified lack of coronary angiography in identifying coronary artery pathology in the setting of myocardial infarction (Reynolds HR, Srichai MB, Iqbal SN, Slater JN, Mancini GB, Feit F, Pena-Sing I, Axel L, Attubato MJ, Yatskar L, Kalhorn RT, Wood DA, Lobach IV, Hochman JS. Circulation 2011 Sep 6). This is just another example of how coronary anatomy and physiology are unique in women and thus the diagnostic approaches may need to be modified in women.	We agree that there is clearly a higher incidence in women with ACS or acute MI without significant obstructive coronary artery disease. This may be as high as 9% in some studies. The report referenced is on 50 women without obstructive CAD and ACS. However, the purpose of this review was to identify the relationship between imaging modalities in women with chest pain and invasive coronary angiography, which despite the stated limitations remains the reference standard for the majority of studies.
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	The onset of chest pain during a stress test, particularly if it corresponds to a perfusion defect on nuclear imaging, or wall motion abnormality on stress echo, despite a normal angiogram, should warrant more aggressive medical management (beta-blockers, calcium channel blockers, nitrates, etc). The Draft Report does not appear to fully address this issue. Indeed, multiple studies have shown that the limitations of coronary angiography, not to mention its risks. Noninvasive tests may be equally predictive or even better predictive of events during follow-up. (From AM, McCully RB, JASE 23:207-14, 2010) and (Sicari R, JACC 41:589-95, 2003).	The reviewer makes a reasonable clinical point; however, the majority of studies do not provide a correlation between the actual symptoms during testing and findings or outcomes. In fact, the two studies reported were not included because they were noncomparator studies. The purpose of this report was not to state that patients with ischemia, microvascular obstruction, or significant spasm should not be treated, but rather to provide information on the diagnostic and available comparative data.

Commentator & Affiliation	Section	Comment	Response
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	In fact, in clinical practice (and as the authors acknowledge) coronary angiography involves risk that is not insignificant, and clearly is not appropriate for all symptomatic patients. There is a large subset of symptomatic women with intermediate risk of heart disease for whom coronary angiography is not indicated based on current professional guidelines. Does the conclusion that there are few comparative studies addressing the capability of NITs to facilitate risk stratification, decision-making and improved outcomes as compared with coronary angiography mean that there is insufficient basis for performing these tests in the clinical environment? In light of the absence of a sufficient volume of comparative studies, do the authors recommend that all patients with suspected coronary artery disease should undergo coronary angiography? A more clinically useful analysis might relax the selection criteria and focus on an evaluation of the available clinical literature regarding whether, and to what extent, each NIT provides useful data in women regarding risk stratification affects decision-making or improves clinical outcomes as compared with stress ECG alone.	We have stated the fact that “comparator” studies have insufficient data. We have stated in the Limitation section that noncomparator studies were not reviewed.
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	It would be remiss if we did not address the issue of exercise capacity and choice of imaging modality. Shaw et al recently found that there was no incremental benefit of exercise MPI compared with standard ETT in symptomatic women. Imaging would therefore appear to play a much more important role in the evaluation of the patient with limited physical capacity. Given the prevalence of obesity in the United States, unfortunately many women are functionally limited by their obesity and associated health problems and routine ETT testing would not apply to them (Shaw, Circulation 2011, 124:1239-1249).	The search was updated in September 2011 and we identified the reference comparing ETT to MPI that was cited (Shaw 2011); it has been included in the KQ 3 section as a comparison of exercise ECG and SPECT (to keep NIT terminology consistent with the rest of the report).
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	In summary, this analysis, while examining a large volume of literature, does not appear to draw clear conclusions about the ideal NIT to use in women with chest pain. Given the current economic climate, cardiac MR and CTA are unlikely to become the initial test in women with chest pain. This leaves ECG stress testing, stress nuclear imaging and stress echocardiography. Stress ECG, for which stress cardiology may be the optimal non-invasive test, appears to be best used in women with an adequate functional capacity. It is helpful in providing prognostic information. It is limited by its limited specificity, in women, particularly post-menopausal women. Nuclear imaging, while sensitive and specific, involves radiation exposure, which is of particulate concern in women. For women with limited exercise capacity, pharmacologic stress echocardiography appears the most viable option.	Noted

Commentator & Affiliation	Section	Comment	Response
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	We do agree with the analysis insofar as it does make the point that further study is needed. Careful attention must be paid to the design and analysis of studies with more emphasis on outcomes and less on the identification of coronary artery stenosis.	Noted
Thomas, James; Wood, Malissa (TEP member); and Mangion, Judy (American Society of Echocardiography)	General	[attachment: American Society of Echocardiography News, Volume 19 Number 2]	The attachment from the American Society of Echocardiography News was reviewed.
Peer Reviewer 2	Executive Summary	The introduction lacks a reference or any details about how the sources of AHA collects their data. Many calculations are based on these data. Other sources on prevalence and incidence are better described and nicely reviewed by Lloyd-Jones. Is the statement from reference 1 independent of age? Later text suggests it is not. The term NIT for noninvasive testing is not familiar to me -- I needed to return to this while reading (but I know it now...).	We have confirmed the prevalence/incidence statistics from references 1 to 3 in the first paragraph of the Introduction. The statement from reference 1 is independent of age.
Peer Reviewer 2	Executive Summary	The ES-2 flow chart is very useful. I was thrilled to see the effort to report the comparisons by sex and age. I liked analyses by good quality studies and all studies with the feel for the numbers presented in the excellent Table ES-3.	Thank you.
Peer Reviewer 2	Executive Summary	I think the term "spectrum bias" is not familiar to many who will want to read this document.. A little help for the reader beginning on pg ES-18.	Spectrum bias (or spectrum effect) is when the performance of a diagnostic test changes between different clinical settings owing to changes in the patient case–mix, thereby affecting the transferability of study results in clinical practice.
Peer Reviewer 2	Executive Summary	pg ES-14 same sentence beginning on line 19 appears twice.	We have removed the duplicate sentence.
Peer Reviewer 4	Executive Summary	A few comments that could be considered. Page 11: line 7. This could read, "…the use of catheter based techniques has superior spatial and temporal resolution when compared to NIT. However, flororoscopy is inherently 2-dimensional while new techniques such as CTA are inherently volumetric."	We agree that fluoroscopy is inherently two-dimensional, but, in general, multiple orthogonal views are obtained at catheterization, which provide reference standard information. Although CTA is volumetric, it is not clear that volumetric analysis was being used as a standard practice in all studies.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 4	Executive Summary	Page 23-24. I believe there should be some careful considerations of "how" CAD is diagnosed, in particular for the newer technologies, CTA and CMR. CT has direct visualization of the coronary arteries and thus this should be straightforward. CMR is more complex. I assume that the analyses relate mostly to cardiac function, but there may be perfusion, and even attempts at imaging coronaries. I am not surprised that the overall numbers are slightly inferior to CT, but there should be some attempt to dissect out which of specific pulse sequences/ methods were used in the MR protocols. The inherent problem with this approach would be to lessen the impact by splitting studies. However, if the pooled numbers are used, maybe add a mention to the fact that there are several different protocols that can be used?	We appreciate the reviewer's point that, in fact, CMR provides a toolbox of techniques that may help provide a higher level of diagnostic accuracy when used well in any given patient. However, this review of the literature did not find one standard set of protocols or dominant diagnostic strategies with CMR; and therefore within the current analysis, CMR is considered with all the techniques used. However, future analysis will need to reconsider this approach.
Peer Reviewer 4	Executive Summary	Page 24. Regarding radiation, there are some obvious differences (e.g. breast tissue) that make effective dose different between women and men (even for the same level of exposure). This is only touched upon, and very lightly, later in the document. I think this will be of at least some interest to the readership. Also, based on my knowledge of the literature, there is no known gender difference with respect to any complication related to contrast (iodine or gadolinium) administration.	We added some sentences in the Results section for KQ 4 that briefly discuss differences in radiosensitivity in reproductive organ tissues. One study we reviewed (Weustink 2009) described long-term complications related to the use of contrast.
Bruening, Wendy	Executive Summary	In my opinion it is too long and unnecessarily detailed for an ES particularly in the description of the methods sections and the inclusion of the study flow diagram and analytical framework figures.	We have used the AHRQ template for executive summaries, which are intended to be standalone documents.
Shaw, Leslee (American Society of Nuclear Cardiology)	Executive Summary	On page ES-2, the authors write: "These limitations are generally considered to be minor when compared with the benefits of the procedure, and coronary angiography is now the standard for clinical care of patients who have chest pain suggestive of CAD." We believe this statement is incorrect given that there are about 1-2 million diagnostic angiograms performed each year and there are well in excess of 10-12 million noninvasive tests performed annually.	We have modified the sentence to state that catheterization is the reference (gold) standard.
Shaw, Leslee (American Society of Nuclear Cardiology)	Executive Summary	On page ES-7, exclusion criteria: we would note that the electron beam CT technology is basically out-dated and this description should be replaced with the title coronary artery calcium scoring. On this same page, the authors write that the setting would be primary care or cardiology clinics. We believe this statement should be clarified as most of the evidence is derived from stress testing laboratories and not clinics.	We have added coronary artery calcium scoring (e.g., electron beam CT) to the exclusion criteria table.

Commentator & Affiliation	Section	Comment	Response
Shaw, Leslee (American Society of Nuclear Cardiology)	Executive Summary	We urge the authors to define the precise handling of duplicate patient series on page ES-10 and include the years of publication.	Duplicate articles found in other literature databases were removed from our project database, with retention of the original reference. The search date for publication is noted in the Executive Summary and the full report.
Shaw, Leslee (American Society of Nuclear Cardiology)	Executive Summary	On page ES-13, the authors write: "The sensitivity of ECHO and SPECT was significantly larger than that of ECG." We believe the term "larger" should be replaced with "greater."	We have replaced "larger" with "greater."
Shaw, Leslee (American Society of Nuclear Cardiology)	Executive Summary	In the discussion on page ES-17, the authors write about the use of pretest probability estimates. It would be helpful to define how these are calculated. Also, in this section, the authors state that "…clinicians faced with patients who have a guideline-defined low pretest probability of CAD (less than 10% chance)-may decide to obtain a noninvasive test, ideally with a high negative predictive value in this population and low risk of adverse events, in order to 'rule out' disease." We are concerned that the use of stress testing in low risk patients is not supported by ACCF / AHA stable ischemic heart disease guidelines or ACC appropriate use criteria.	We agree and have changed the text to read "low-to-intermediate pretest probability." Pretest probability is calculated with AHA-recommended methods, such as Diamond-Forrester or CASS study criteria (where pain type, age, and gender were powerful predictors of the likelihood of CAD) or the Duke database criteria (which also includes risk factors).
Shaw, Leslee (American Society of Nuclear Cardiology)	Executive Summary	Further, we believe the ensuing paragraph is highly biased and inaccurate and suggest that it be re-written. a. The authors write: "First, women by definition are thought to be at lower pretest probability of CAD when evaluated in comparison with men of the same age." This is far from being consistently accurate. In many cases, women have a higher risk factor burden and more co-morbidity than men.	We agree with the reviewer, but the current recommended method by the AHA stable angina guidelines shows that women compared with men of the same age are considered lower risk (although not always the case with comorbidities/risk factors).
Shaw, Leslee (American Society of Nuclear Cardiology)	Executive Summary	The authors also write: "Second, women may be uniquely susceptible to some of the adverse effects of testing, may have poor test performance, involving radiation or have higher rates of complications from invasive arterial access." As noted above, there are no observed "adverse effects" of radiation. We urge the authors to choose their words carefully with regard to the available evidence on ionizing radiation exposure. We reiterate that the projected risks are generally estimated to be higher due to the greater radiosensitivity of reproductive organs to ionizing radiation.	We have modified the sentence to remove the focus on radiation.
Shaw, Leslee (American Society of Nuclear Cardiology)	Executive Summary	The authors also state: "Third, because of body and chest morphology due to body shape, women may not obtain the same test performance that men do from noninvasive testing." The authors should be more specific here as it is unclear exactly what they are referring to.	We have modified the sentence to replace "body and chest morphology" to "body shape."

Commentator & Affiliation	Section	Comment	Response
Shaw, Leslee (American Society of Nuclear Cardiology)	Executive Summary	On page ES-19, the authors note the use of MR angiography. We are concerned that this modality is employed in clinical practice and suggest deleting this reference. We also suggest eliminating the entire paragraph on functional vs. anatomic imaging as it is off-base for this report. Moreover, we are concerned that such an expert discussion may be beyond the grasp of many readers of this document.	Other reviewers felt that MR angiography should be included and that the distinction between function and anatomic imaging was a strength.
Peer Reviewer 1	Introduction	Introduction: This section was well written and justified the need for this comparative effectiveness review. The phrase "chest pain syndrome in women" should be defined in this introduction.	We have changed the language “chest pain syndrome” to “suspected CAD” or “symptoms suspicious for CAD,” which includes both angina and nonanginal chest pain and is the term used in the majority of the studies to describe symptomatic women undergoing an evaluation.
Peer Reviewer 3	Introduction	Introduction: The introduction is succinct, and describes the growing global problem of heart disease in women, including the specific challenges in diagnosis and prognostication of coronary disease in women. There is a brief description of each of the noninvasive methods studied, along with limitations and strengths of each. Finally, the objectives of the review are spelled out towards the end of the introduction section. A suggestion here is to emphasize the difficulty in assessing pre-test likelihood of coronary disease in general and women in particular. Since choice of testing depends on pre-test likelihood assessment, it is essential to describe limitations in current algorithms and also stress this in the "Future Research" section. This reviewer would also like an explanation of why calcium scoring cardiac CT was not included in the analysis. If this is because CAC is primarily used in asymptomatic subjects, this must be stated. However, CAC has been used in the acute chest pain setting with multiple published reports. Because this is again tied to pre-test likelihood assessment, it is important for the reader to understand where CAC stands in the spectrum of non-invasive testing.	We agree and have added language that the estimation of pretest probability is central to determination of and selection of noninvasive tests for CAD. Additionally, we have added a statement that coronary artery calcium scoring is not routinely used in symptomatic patients and therefore was not considered for this report.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 6	Introduction	Introduction: The background section of the introduction provides some detail on the indication, procedure, and risks of coronary angiography as the gold standard for diagnosis. The role of NIT almost appears as an alternative when contraindications or risk of complication are high, as opposed to having an inherent value and role apart from angiography. Perhaps some minor tweaking of this section could really highlight the niche of NITs as the focus of this review. Another potential enhancement could be a bit more technical information on the details of ischemia assessment with the different modalities of NIT. For example, the section on Echo modalities details the parameters that are analyzed and the stress modalities used - exercise and pharmacologic - in combination with imaging. However, the section on MPI modalities does not detail the stress modalities utilized and how MPI is interpreted to diagnose ischemia. Additionally, the emerging modalities category explains the concept of functional versus anatomic assessment. However, describing this concept in an expanded description of the role of NITs above the individual descriptions of each individual NIT may be helpful. Again, I think it may be important to provide some detail on how obstructive disease or ischemia is assessed with each modality, differentiating exercise versus pharmacologic testing and commenting on functional versus anatomic assessment.	More information about the criteria for a positive test for each NIT modality has been added to the report. The heading “Emerging Modalities” has been replaced with “Anatomic Modalities” to distinguish between the functional and anatomic assessment.
Shaw, Leslee (American Society of Nuclear Cardiology)	Introduction	On page 1 the introduction presents invasive angiography as a preferred option noting the ability to combine treatment and diagnosis at the same procedure. We believe this line of thinking is misguided considering the focus of ischemia-guided medical and invasive therapeutic management.	We have noted the limitations of invasive angiography in the Introduction.
Shaw, Leslee (American Society of Nuclear Cardiology)	Introduction	On page 2 ECG modalities are addressed. We suggest deleting reference to multifunction cardiogram as this is not a part of routine clinical practice.	The multifunction cardiogram modality was added in response to a public comment by Premier Heart.
Shaw, Leslee (American Society of Nuclear Cardiology)	Introduction	On page 3 within the Echo section, we suggest noting that vasodilator Echo is rarely performed in the US. The notation of a lower sensitivity in women with stress Echo should be framed within the research that has shown a reduced sensitivity in patients with less extensive CAD (i.e., single vessel CAD). It would be important to note the limitations in obese women and those who smoke of reduced left ventricular opacification.	We acknowledge your comments but were unable to find references that support these statements, and therefore did not add this information to the Introduction.

Commentator & Affiliation	Section	Comment	Response
Shaw, Leslee (American Society of Nuclear Cardiology)	Introduction	On page 4 in the "Uncertainties" section we have several questions/comments for the authors. a. It is unclear what "different response to exercise" means; b. We suggest removing mitral valve prolapsed from the non-obstructive CAD example.; c. It is unclear what lower peak exercise values the reviewers referring to. Exercise duration is also not well defined; d. It is unclear what is meant by an inappropriate catecholamine response; e. We suggest combining non-obstructive and single vessel CAD to state a lower prevalence of CAD; f. It is unclear what anatomic differences affect stress test results; and g. Instead of stating breast attenuation artifacts, we suggest stating reduced accuracy in obese women and/or those with large breasts.	The details for the definitions of these terms can be found in references 3 and 4. We do not believe that repeating the detailed definitions (and supporting evidence) from those references is necessary for making the point that there are many factors that may affect diagnostic accuracy in women.
Shaw, Leslee (American Society of Nuclear Cardiology)	Introduction	Finally, the ensuing paragraph on order of diagnostic testing fails to note that several guidelines have a recommended "work-up" strategy. We suggest deleting the sentences on use of angiography in the post-ACS patient. As well, in the "Relevance" section, there are recommendations for the use of non-invasive testing including the use of the exercise ECG in intermediate risk women who are functionally capable of maximal stress testing and who have a normal rest ECG. For intermediate-high likelihood women with an abnormal ECG, either Echo or nuclear imaging is recommended. For women incapable of exercising, either pharmacologic stress Echo or nuclear imaging is recommended. More details of this diagnostic work-up can be located in the ACC / AHA stable angina guidelines or in the 2005 AHA statement on cardiac imaging in women (Mieres Circulation 2005).	We agree that the diagnostic workup strategy is recommended with conflicting guideline recommendations. The noted AHA/ACC stable angina guidelines recommend exercise ECG in patients with the ability to exercise and a normal ECG. The stress ECHO and nuclear guidelines recommend stress imaging**.**
Peer Reviewer 1	Methods	Methods: The inclusion/exclusion criteria were explicitly stated, justified, and logical. The search strategies were explained in very specific detail and appear exhaustive. On ES-8 and on page 8, the authors state the aabstraction templates were "pilot tested to ensure consistency and reproducibility" but no further information was provided. Were refinements necessary? If so, what? Who pilot tested and how? Were the persons abstracting for the pilot the same as for this review? What was actual inter rater reliability? Study quality assessment was clearly defined. Statistical methods were clearly stated, presented, and appropriate.	The data abstraction forms were piloted by two members of the study team, and refinements to clarify the wording of the questions and collection of the data (numeric vs. text, adding "other" and "please specify" as a potential choice) were added after the first week of data abstraction. The investigators who piloted the forms were also the main data abstractors. The forms were overread by two other members of the team, and any changes to the form were noted in a comment field. The overall interrater reliability was 0.76.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 3	Methods	Methods: The methods of performing this review are exhaustive, thoughtful and detailed. The authors performed a literature search of relevant studies over the last decade in the English language, with two independent reviewers reading each fulltext article to determine inclusion into the analysis. Furthermore, two researchers abstracted all data from selected studies using specifically created forms. All studies were graded as good, fair or poor based on AHRQ's Methods Guide and from elements of QUADAS. Statistical analyses were robust and adequate for this report.	Thank you.
Peer Reviewer 6	Methods	Methods: The methods nicely report on the strategies used to search the literature for articles and reports related to this comparative effectiveness topic. Additionally, the figures and tables are quite helpful in elaborating on the approach. I had a few comments.	Thank you.
Peer Reviewer 6	Methods	Minor comments - On p. 7 line 1, the abbreviation EPC is used for the first time without explanation. On that same line, the authors refer to the recommended methodology, but there is not a reference or description of this methodology.	We have added the definition of Evidence-based Practice Center and referenced AHRQ's methodology document as well as describing the methodology throughout the Methods section.
Peer Reviewer 6	Methods	The analytic framework on p. 8 is a nice display of how the key questions fit into the algorithm of patient care and decision-making. However, a figure legend or perhaps footnotes may help to clarify a few things. One question that arose for me is what is meant by prevalence in the box labeled "Diagnosis." The other bullets refer to specific properties of the test, whereas prevalence is a population measure which may influence these test parameters. Perhaps the list represents the elements that would be reported for this diagnostic accuracy question. A figure legend could frame the intention of the elements displayed. Under the box entitled "Predictors" the term stress modality may not be as inclusive as the authors intended. For example CTA may not also have a "stress" component except in research protocols. Perhaps "NIT modality" or "testing modality" may better reflect what it is meant. Lastly, in the box labeled "Clinical outcomes" no obstructive CAD or non-coronary etiology is a possible outcome that is not represented.	The list does include the elements that would be reported for the diagnostic accuracy question. Given the limited space in the analytic framework figure, we list only the general categories rather than the specific circumstances for each NIT modality. Also, the "no treatment needed" and/or "medical management" options in the Treatment box is meant to represent the "no obstructive CAD" outcome.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 6	Methods	For inclusion and exclusion criteria, Table 1 beginning on p. 10 and Figure 2 on p. 16 are both necessary to understand the exclusion criteria and the number excluded for each criteria. A minor suggestion is to include the number excluded for each criteria in the Table as well to minimize the back and forth that the reader has to do between the two. Were excluded articles evaluated in some way by the authors in order to determine if there was any meaningful information being excluded in articles that were systematically screened out? A minor point in Figure 2, the final box with 101 articles may also need a footnote to represent that the total across the key questions may exceed the number in the box as an article could be used to address more than one key question.	We have reviewed this comment and decided to retain Table 1 and Figure 2 in their current format. Two investigators independently reviewed articles and had to agree on the reason(s) for exclusion, so we do not think that any meaningful information was being systematically screened out. We have added a second note in Figure 2 to state that an article could have been used to address more than one key question.
Peer Reviewer 6	Methods	The definitions for the diagnostic criteria for the outcome measures are clear and well defined, mostly for KQ1 and KQ2 which summarized accuracy data and reported statistical estimations. KQ3 and KQ4 also had clear definitions and outcomes measures which were reported in a descriptive format without additional statistical analysis.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Bruening, Wendy	Methods	I do not understand what method was used to compare the different NITs. The methodology section has no description of the method used. There don't appear to be any studies directly comparing the different NITs. Therefore the method used must be an indirect method. However it appears to involve statistical testing. I am not aware of any well-accepted, well-studied method of indirectly comparing the accuracy of diagnostic testing modalities. The only published, somewhat accepted method I know of is the conditional relative odds ratio (Suzuki et al.) which is clearly not the method used here, because the conditional relative odds ratio doesn't analyze sensitivity and specificity separately. So it would be nice if the authors described their methods for performing the indirect comparison and provided some kind of validation of the method used. The provided forest plots don't support the claims of statistically significant differences between NITs because the confidence intervals appear to overlap quite significantly. More attention to how "homogeneity" was determined would also be nice. The authors mention concerns about homogeneity were used when deciding when to perform meta-analyes, but don't explain how homogeneity was assessed. Also many of the SROCs look rather heterogeneous, but no statistics such as I^2 are presented.	We described the method in the Data Synthesis section. We compared performance of testing modalities within the random effects framework used to obtain summary values of sensitivity and specificity for each modality separately. Namely, we considered a data set comprising all modalities with a covariate defining modality (i.e., categorical variable with 5 categories). P-values arise from testing for presence of interactions. The overall p-value indicates that there is at least one significant difference, and the forest plot with summary values by modality indeed shows that at least one 95% CI does not overlap with one of the others. Potential heterogeneity between studies is reflected though the summary confidence intervals (CI) and regions (CR) obtained from a random effects approach used in this report for all the analyses. These CIs and CRs are larger than those obtained from a fixed effects analysis. We do not present fixed effects results in the report.
Peer Reviewer 1	Results	Results: The presentation of results by key questions and key points, etc were easy to follow. I particularly found the key points section, such as that on page 18, to be extremely helpful. Presenting the sensitivity, specificity, etc by quality of study, such as in Table 3 (starting on page 40), was especially informative and clinically relevant. The Figures, tables, appendices were clearly presented, informative, and exhaustive. I did not identify any studies that needed to be included that were published in the specified timeframe.	Thank you.
Peer Reviewer 2	Results	Should I assume that all ECGs are GXT and not stress ECGS. This may be in a table but I missed it in the text. Are all the ECGs in Figure 6 GXTS or do they include the newer method you describe?	Two of the ECG studies employed multifunction cardiogram and two studies were pharmacologic stress; all the rest were exercise testing. This detail has been added to the ECG results section.
Peer Reviewer 2	Results	Could the QUADAS be in a box for easy access; these criteria are not memorized by the rest of us.	The QUADAS questions are outlined in Appendix B in the list of data abstraction elements.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 2	Results	Important that for ECGs, despite relatively small numbers, most of the 95% CI are relatively similar whether or not the point estimate is to the left or right of the pooled sensitivity or specificity statistic. The CI are even narrower in Figure 10 with only 13 good quality studies.	We agree.
Peer Reviewer 2	Results	I was not familiar with a "SROC", but I like the ability to show sensitivity and specificity together with boundaries this way. I did not understand the methodology. I do think more information on the method and on added value of this analysis in the same place the results are shown would be helpful.	The SROC method is described in the Data Synthesis section; The Leeflang 2008 reference contains more information.
Peer Reviewer 2	Results	I was disappointed in the limited data on comparative costs of the actual tests, limited data on untoward events caused by the method, and limited consideration of the external validity which may be very low based on the source of most of the studies. It is not clear whether these studies used data from specialists in referral hospitals who may be better at the tests or their interpretation?	The comparative cost of these tests is outside the scope of our report. We agree, however, that there are limited data on untoward events in these studies. Many of these studies were conducted at academic setting; whether study interpretation is better or the same at academic or nonacademic settings is unknown. This could depend on the volume of studies conducted at the site, and this information is not available in the published literature.
Peer Reviewer 2	Results	I would have liked more early discussion of efficacy (trial) before you got into the real world of efficiency -- that gave this reader more of a feeling of results in the best possible circumstances.	Trials for efficacy (or RCTs) with coronary angiography as the gold standard are rarely performed in the NIT literature; therefore, the focus is on the effectiveness from primarily observational studies
Peer Reviewer 2	Results	I wondered why there was no mention (or did I miss it) of intravascular ultrasound, which I believe detects lesions not detected by your coronary artery angiogram gold standard. The WISE studies concluding no obstructive disease in their women have been criticized for no IVUS, and I understand they have now been funded to do IVUS. If this is not included because there are too few comparative studies meetiing your criteria, it still should have been discussed. If IVUS is not widely used then you should say so.	IVUS is performed only during invasive cardiac catheterization and is not considered a noninvasive test.(NIT)
Peer Reviewer 2	Results	KQ4 had an appropriate concern about radiation, but the number of methods permitting comparison was shocking. I was not aware of the excess cancer risk, and wanted to know what cancers these are (i.e., in the direct line of radiation or not). Other radiation-related methods such as CT angiography require medication (e.g. to slow heart rate) and carry some risk.	The estimates for excess cancer risk were derived from a calculation using a risk factor of 0.05 per sievert. The location and type of cancers that might result from various levels of radiation exposure were not discussed in either the paper that described the risk estimation methods or the paper that used these methods to estimate excess age- and gender-specific cancer risk associated with CT angiography.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 2	Results	How many test procedures cannot be completed? I assume this differs by test method. Some of these tests identify other "abnormalities" leading to other procedures, most of which were not of clinical significance (incidentalomas). Not much said about this additional work-up risk.	Withdrawals and incomplete tests were rarely reported in these studies; likewise for reporting of incidentalomas.
Peer Reviewer 3	Results	Results: Diagnostic accuracy of each non-invasive method was determined by comparison to invasive coronary angiography. A summary ROC curve was generated to depict sensitivity and specificity of each modality, separately in those without known CAD and in "mixed" populations of known/no known CAD. Finally, a comparative analysis was performed for the diagnostic accuracy of the modalities in women compared to men. Coronary CT angiography had the highest sensitivity and stress echocardiography the highest specificity. Compared to men, the specificity of stress echocardiography was lower in women with statistical significance. No other gender differences were noted in diagnostic accuracy. There was insufficient data to determine the predictors of diagnostic accuracy. Although one of the key objectives, there were only 7 studies that provided any evidence on risk stratification and prognosis, only 2 on clinical decision-making and 3 providing data on clinical outcomes, limiting this aspect of this report significantly. Another key objective identified was the safety of each modality. Once again, only a handful of studies (4 ECG, 5 echo, 3 MPI, 1 CMR and 4 CTA) were available, thus significantly limiting this aspect of this report as well.	We agree.
Peer Reviewer 6	Results	Results: The detail in the results section is very appropriate for each key question described. The key messages are explicit in the format in which they are presented. The characteristics of the studies are well descirbed by the summary tables.	Thank you.
Peer Reviewer 6	Results	For KQ1, the summary tables are well designed and the SROC curves quite nicely demonstrate the trade-off between sensitivity and specificity.	Thank you.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 6	Results	For KQ2, I believe that the tables report test characteristics such as sensitivity, specificity, PPV, NPV, LR in the subgroup identified (eg., Table 10, line 16 is the data for the subgroup of Hispanic women in this study). However, it was unclear from the tables how the test performed in the reference subgroup. For example in the study with menopausal status as a predictor, do we know how this compares to non-menopausal women? Or in the study with race/ethnicity as a predictor, do we know how these statistics compare to Caucasian women? If not available in the studies, a footnote indicating such may be helpful. The tables on heart size and other predictors do provide the statistics on the referent group.	All women were postmenopausal (i.e., no reference group of premenopausal women). Similarly, the paper on minorities did not have a Caucasian comparison group. Footnotes have been added to these tables to show the lack of a reference group.
Peer Reviewer 6	Results	For KQ3 and KQ4, the number of studies available were so few, that I wondered if a broader inclusion search strategy might provide some useful information for directing future research. For example, particularly for KQ4, would there have been information gleaned from studies that may have been categorized as "non comparator studies." Perhps these studies would report events that could be compared across studies of differing modalities including the gold standard (coronary angiography).	We considered a broader inclusion of noncomparator studies for KQ 3 and KQ 4; however, the selection bias, referral bias, etc., from including these types of studies would have made the indirect comparisons for the KQ 3 and KQ4 outcomes problematic for quantitative analysis.
Peer Reviewer 1	Summary and Discussion	Discussion/ Conclusion: Table 7 in the summary section, page 64 & 65 and indeed all tables in the summary section are very useful and provide an excellent summary of the findings. Areas where insufficient data were present were clearly stated, and these areas were then discussed in the future reseaach section. Areas requiring significant confirmation of results were clearly identified. In the discussion and limitation sections, I highly recommend including comments related to chest pain syndrome in women and the need for a more precise description of potential symptoms in women. The women who present without chest pain are often the most difficult to diagnosis with CHD and these diagnostic NIT modialities are not used in them, primarily because of the lack of chest pain. I was very pleased that the authors emphasized the need for CHD studies of sufficient size and represenation of women so that meaningful analyzes may be conducted between sexes. I woulds also recommend further research looking for possible racial differences.	The term “chest pain syndrome” has been replaced with “symptoms suspicious for CAD” or “suspected CAD.” We have added that future research should also look at race differences.
Peer Reviewer 2	Summary and Discussion	The discussion is really excellent with regard to pre-test probablility and also sex differences. Is the term chest morphology (ES-18) a euphymism for breasts?	We have replaced “body and chest morphology” with “body shape.”

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 3	Summary and Discussion	Discussion/ Conclusion: The discussion section could be lengthened a bit to specifically discuss issues related to pre-test likelihood assessment, availability of tests and influence of accuracy, and the role of emerging imaging modalities. The issue of microvascular disease in women could be expanded upon - for example, what is the prevalence of this? How does this influence comparative studies of noninvasive testing and invasive coronary angiography? Would modalities that combine anatomic and functional imaging (such as SPECT/CT, PET/CT or CT perfusion) be of specific use in these women? Why were these studies not included? This reviewer would like at least a marginal discussion on using invasive coronary angiography and not intravascular ultrasound or fractional flow reserve as the "gold standard".	We have rewritten the pretest discussion. We have noted the limitations regarding microvascular obstruction—the incidence of which is hard to determine since there is no clear diagnostic test used all the time. Also we have noted that 9% of ACS in women may not be obstructive CAD. The combinations of modalities listed by the reviewer were not included unless they were in comparison to coronary angiography; plus these studies continue to evolve and therefore are less represented in the literature. Traditionally, invasive coronary angiography is the reference standard, and IVUS or FFR require further intravessel assessment. This is not without risk but is increasing in use for equivocal lesions.
Peer Reviewer 5	Summary and Discussion	1) Discussion page 100 line 10 - delete by definition, change to women may be thought to be at...	We have made this change.
Peer Reviewer 5	Summary and Discussion	2) It would be helpful to expand limitations discussion to include single sex studies limit the ability to evaluate gender differences.	We have expanded the limitation section.
Peer Reviewer 5	Summary and Discussion	3) Is it appropriate to address the issue the performance of test may be location/operator dependent and the results from clinical research trials may not uniformly apply to practice as a bit of a disclaimer. Since the review was so labor intensive by experts, giving more definitive guidance on limitations and recommendations would be helpful.	We agree and have added information about this.

Commentator & Affiliation	Section	Comment	Response
Peer Reviewer 6	Summary and Discussion	Discussion/ Conclusion: The implications and major findings are clearly stated. The future research section highlights some important concepts. The section on RCTs comparing NIT modalities cites two trials that are already funded and ongoing. Perhaps highlighting that there are remaining gaps in the science that are not being answered by these trials could bolster future investigation. Additional questions across testing modalities may include women at various risk profiles, in varying settings (e.g., outpatient, inpatient, emergency department), for differing symptamatology, and timing in presentation(s) for example. The other sections in this Future Research aspect of the report nicely refer to characteristics that would be important in any future studies.	We agree and will add these statements.
Bruening, Wendy	Summary and Discussion	A limitation of the evidence base that isn't discussed in detail is the possibility of spectrum bias- presumably the intent/ future use of NITs is as a triage test to decide who should go on to invasive angiography testing. However, all of the included studies only included patients who had already been referred for invasive angiography testing. Therefore it is very possible that the studied patients have higher pre-test probabilities of CAD than the actual patient population of interest. This should be discussed and possible impact on the applicability of the review's findings to clinical practice should also be discussed.	There is a discussion about spectrum bias in the third paragraph of the Discussion section of the report. Not all studies included patients who had already been referred for invasive angiography. Some of the studies included another NIT modality, as well as invasive angiography, as a method for evaluating diagnostic accuracy.
Peer Reviewer 3	Future Research	Future Research: This section is comprehensive. This reviewer would like to see this expanded to include other, novel imaging tools and potential advantages in women. Layered testing as a focus for future research must be emphasized along with appropriateness of testing. Gender-specific issues such as symptom and test-related office visits, downstream testing that can be prolonged, use of medications, etc must be addressed as research criteria.	We agree and have already discussed layered testing in the Discussion section. We have added text in the Future Research section for an assessment of existing appropriate use criteria for the different NIT modalities and the need for an evaluation of diagnostic algorithms to ensure strategies for improving outcomes are tested.