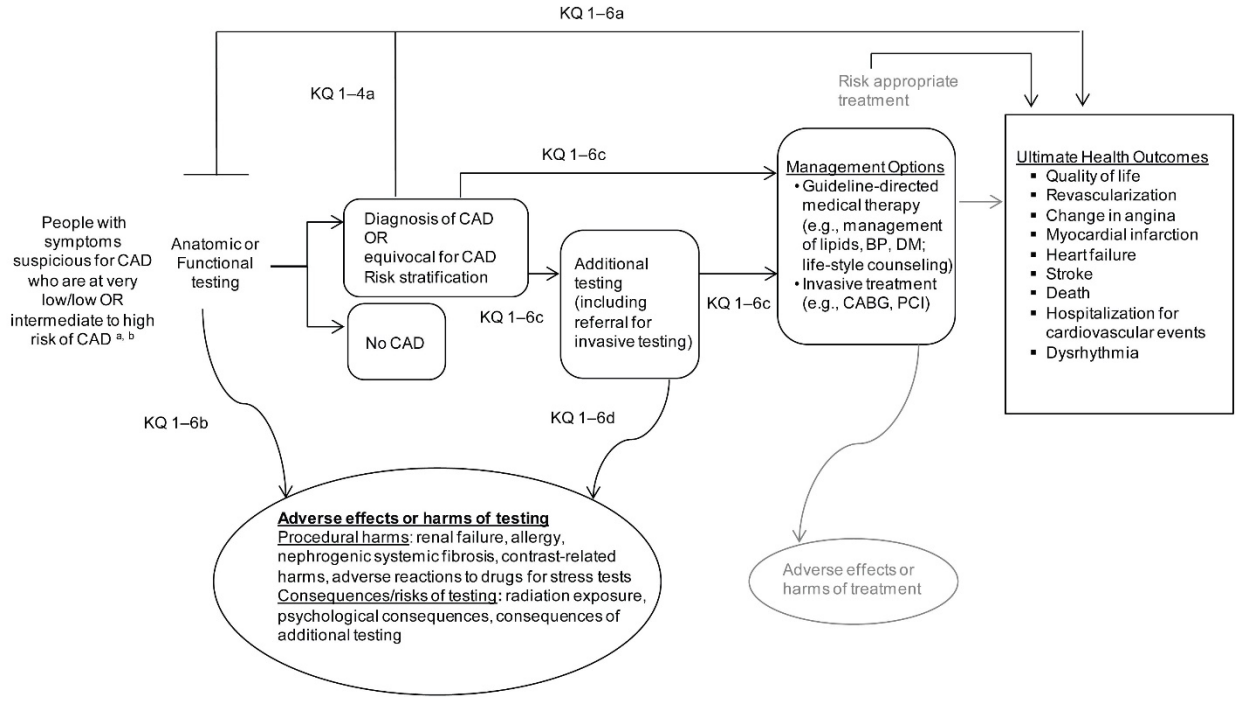
III. Analytic Framework

Figure 2. Analytic framework for noninvasive testing for coronary artery disease



[a] People at very low or low risk will be evaluated separately from those at intermediate to high risk as possible.
[b] KQ 1–6e: Potential modifiers related to differential efficacy and/or safety include patient factors (e.g., age, sex), comorbidities, and ability to exercise.
BP = blood pressure; CABG = coronary artery bypass graft; CAD = coronary artery disease; DM = diabetes mellitus; KQ = Key Question; PCI = percutaneous coronary intervention.

IV. Methods

Input from the Technical Expert Panel (TEP) affirmed that a focus on the associations between testing and clinical outcomes would be of primary interest but there is still a need to provide some information on traditional test parameters as a foundation. Thus, to set the stage for the review's Key Questions, which are focused on clinical decisionmaking and clinical outcomes, **contextual** information on the following relevant topics of will be provided in the background of the report.

- Description of how pre-test probability/risk of CAD is determined based on the triad of patient presentation and physical exam, family history of CAD, and findings on resting ECG in usual clinical practice
- The ability of ICA to predict primary clinical health outcomes (e.g., avoiding myocardial infarction)
- Brief discussion of limitations of ICA (thresholds, reliability, etc.) as a reference standard
- Brief summary of the diagnostic accuracy, in terms of traditional test performance measures (e.g., sensitivity, specificity) of each of the following noninvasive tests based on the highest quality systematic reviews available in symptomatic patients with suspected CAD compared with the historical gold standard of ICA, for the tests listed below. (As ICA with FFR is generally used in people with known CAD, where there is literature using it as a referent in the population of interest it will also be described.)
    - Anatomic tests: coronary calcium scoring via EBCT or MDCT, and CCTA
    - Functional tests (including exercise, vasodilator, and/or dobutamine as stressor where appropriate): exercise electrocardiogram without imaging, exercise/pharmacologic echocardiography (with or without myocardial echo contrast), exercise/pharmacologic nuclear cardiology studies (including SPECT and PET), CT perfusion
- Brief summary of the diagnostic accuracy of noninvasive tests compared with each other, in terms of traditional test performance measures (e.g., sensitivity, specificity) of each of the noninvasive tests, based on logical comparisons of tests to each other and with ICA, with or without FFR as appropriate
- Overview of the perceived role(s) of each of the noninvasive diagnostic tests (e.g., to triage, replace, or add on to another test) including:
    - General characteristics of each test and how they are usually employed in clinical settings.
    - For stress testing, provide an overview of when exercises can or cannot be used and general information regarding differences in these populations
- An overview of treatment efficacy in people with stable CAD will be provided based on the concept that an ideal test is one that is safe, sensitive, and specific and for which efficacious treatment is available for test-positive people. Information may include brief discussion of testing parameters and results in relationship to clinical decisionmaking and thresholds for considering various treatment options. Information may include how treatment(s) compare with no treatment to provide a foundation for interpreting trials in which treatments are compared and information is provided on the efficacy of specific treatments.

A. Criteria for Inclusion/Exclusion of Studies in the Review

The criteria for inclusion and exclusion of studies for the systematic review will be based on the Key Questions and are described in the previous PICOTS section and Table 1. Below are additional details on the scope of this project:

Study Designs: Systematic reviews will be used if they address a Key Question, include studies that meet the PICOTS as defined above, and are assessed as being at low risk of bias, according to the AMSTAR (A Measurement Tool to Assess Systematic Reviews) quality assessment tool.[15, 16] If multiple systematic reviews are relevant and low risk of bias, we will focus on the findings from the most recent reviews and evaluate areas of consistency and inconsistency across the reviews.[17] If systematic reviews are included, we will update findings with any new primary studies identified in our searches and assess strength of evidence based on the totality of evidence. Randomized controlled trials (RCTs) as well as observational studies will be considered and those included will be critically appraised. Prospective studies directly comparing interventions with comparators based on established diagnostic criteria will be sought. Retrospective studies will be considered if there are insufficient prospective studies and they are at low risk of bias. Studies of prognosis and decisionmaking will be included testing results are reported in relation to clinical outcomes and if there is appropriate control for confounding. Studies of predictive accuracy will be considered if there are inadequate comparative studies and if data following testing are available for untreated people. Studies using standard of care as a comparator will be excluded if components for standard of care are not explicitly delineated or defined.

Non-English Language Studies: We will restrict inclusion to English language articles, given the large volume of literature written in English on this topic. We will keep track of studies not written in English that would otherwise meet inclusion criteria to provide insight regarding possible language bias

B. Searching for the Evidence: Literature Search Strategies for Identification of Relevant Studies to Answer the Key Questions

Publication Date Range:
Searches will be conducted without restriction on publication date.

Literature searches will be updated while the draft report is posted for public comment and peer review to capture any new publications. Literature identified during the updated search will be assessed by following the same process of dual review as all other studies considered for inclusion in the report. If any pertinent new literature is identified for inclusion in the report, it will be incorporated before the final submission of the report.

Literature Databases: Ovid MEDLINE, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, and International Network of Agencies for Health Technology Assessment's (INAHTA) Center for Reviews and Dissemination (CRD)-HTA database will be searched to capture both published and gray literature.

Scientific Information Packets: Scientific Information Packets (SIPs) will be solicited by the Scientific Resource Center via the AHRQ Web site and via direct mailings to manufacturers.

Hand Searching: Reference lists of included articles will also be reviewed for includable literature.

Process for Selecting Studies: Pre-established criteria will be used to determine eligibility for inclusion and exclusion of abstracts in accordance with the Agency for Healthcare Research and Quality (AHRQ) *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*.[15] To ensure accuracy, all excluded abstracts will be dual reviewed. All citations deemed appropriate for inclusion by at least one of the reviewers will be retrieved. Each full-text article will be independently reviewed for eligibility by two team members, including any articles suggested by peer reviewers or that arise from the public posting process. Any disagreements will be resolved by consensus.

C. Data Abstraction and Data Management

After studies are selected for inclusion, data will be abstracted into categories that include but are not limited to: study design, year, setting, country, sample size, eligibility criteria, population and clinical characteristics, diagnostic criteria and thresholds, setting and rationale for testing/intended role of the test and results relevant to each part of the Key Question as outlined in the previous PICOTS section and Appendix. Information relevant for assessing applicability will be abstracted, including the number of patients receiving testing, characteristics of the population (including CAD prevalence, comorbid conditions), factors that may affect test performance (e.g., presence of arrhythmias), and care setting (both the characteristics of the health care organization and practice patterns). All extracted study data will be verified for accuracy and completeness by a second team member. A record of studies excluded at the full-text level with reasons for exclusion will be maintained.

D. Assessment of Methodological Risk of Bias of Individual Studies

Predefined criteria will be used to assess the quality of included studies. We will focus on studies with the least potential for bias and the fewest limitations. We will assess study limitations using instruments designed to address issues particularly relevant to medical testing studies, such as Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) and Standards for Reporting of Diagnostic Accuracy (STARD) and standardized application of such criteria as described in the

AHRQ *Methods Guide for Medical Test Reviews*.[18] For studies of prognosis and clinical management, risk of bias assessment will be based on pre-defined criteria using clearly defined templates for the criteria as appropriate, following guidance from the AHRQ *Methods Guide for Effectiveness and Comparative Effectiveness Reviews*.[15] Randomized controlled trials (RCTs) will be assessed based on appropriate criteria and methods established in the *Cochrane Handbook for Systematic Reviews of Interventions*.[19] These criteria and methods will be used in concordance with the approach recommended in the chapter, Assessing the Risk of Bias of Individual Studies When Comparing Medical Interventions.[20] Studies will be rated as being "good," "fair," or "poor" quality.

Studies rated "good" are considered to have the least risk of bias, and their results are considered valid. Good-quality studies include clear descriptions of the population, setting, interventions, and comparison groups; a valid method for allocation of patients to treatment; low dropout rates and clear reporting of dropouts; appropriate means for preventing bias; and appropriate measurement of outcomes.

Studies rated "fair" are susceptible to some bias, though not enough to invalidate the results. These studies may not meet all the criteria for a rating of good quality, but no flaw is likely to cause major bias. The study may be missing information, making it difficult to assess limitations and potential problems. The fair-quality category is broad, and studies with this rating will vary in their strengths and weaknesses. The results of some fair-quality studies are likely to be valid, while others may be only possibly valid.

Studies rated "poor" have significant flaws that imply biases of various types that may invalidate the results. They have a serious or "fatal" flaw in design, analysis, or reporting; large amounts of missing information; discrepancies in reporting; or serious problems in the delivery of the intervention. The results of these studies are at least as likely to reflect flaws in the study design as the true difference between the compared interventions. Studies rated as being poor in quality a priori were not excluded, but considered to be less reliable than higher quality studies when synthesizing the evidence, particularly if discrepancies between studies are present.

Each study evaluated will be dual-reviewed for quality by two team members. Any disagreements will be resolved by discussion and consensus.

E. Data Synthesis

We will construct evidence tables identifying the study characteristics (as discussed above), results of interest, and quality ratings for all included studies, and summary tables to highlight the main findings. We will review and highlight studies by using a hierarchy-of-evidence approach, where the best evidence is the focus of our synthesis for each key question. Studies with the least risk of bias will be summarized separately and compared with summarized results from poorer quality studies. In general, prospective studies which directly compare tests of interest in an appropriate

spectrum of patients using validated diagnostic criteria that provide details regarding measurement of pertinent outcomes as well as consideration of and control for bias and confounding may have the least potential for bias. Retrospective studies will be considered if there is low risk of bias. In the evidence tables, we will include relevant studies from included systematic reviews as appropriate.

The intended focus is on clinical outcomes and clinical pathways following the first test performed as result of initial risk assessment (which includes clinical presentation and physical exam, family history of CAD and findings on resting ECG). Harms related to subsequent testing will be evaluated. The specific tests to be covered will be restricted to those identified as widely available and most clinically applicable and established. We will focus on "decision-relevant" outcomes, including the consequences of testing and impact on clinical outcomes such as myocardial infarction and need for revascularization. Consequences of testing may also include use of downstream testing, anxiety related to false positives, and incidental findings. Decision diagrams may facilitate conceptualization of this and augment the current analytic framework. Heterogeneity in study quality is expected and evaluation of the influence of bias will be considered. The categories of very low and low risk of CAD will be combined and analyzed separately from those at intermediate to high CAD risk.

Data will be qualitatively summarized in summary tables or figures. Interpretation of the results will be provided descriptively.

Meta-analyses will be conducted to summarize data and obtain more precise estimates on outcomes for which studies are homogeneous enough to provide a meaningful combined estimate. The feasibility of a quantitative synthesis will depend on the number and completeness of reported outcomes and a lack of heterogeneity among the reported results. To determine whether meta-analysis could be meaningfully performed, we will consider the quality of the studies and the heterogeneity among studies in design, patient population, interventions, and outcomes, and may conduct sensitivity analyses. The key questions are designed to assess the comparative effectiveness and harms of various noninvasive tests for coronary artery disease. Meta-regression may be conducted to explore statistical heterogeneity using additional variables on methodological or other characteristics (e.g., quality factors, diagnostic thresholds or criteria, outcome definitions and ascertainment) if there are sufficient numbers of studies.

Results will be presented as structured by the key questions, and any prioritized outcomes will be presented first.

F. Grading the Strength of Evidence (SOE) for Individual Comparisons and Outcomes

The strength of evidence for each key question will be initially assessed by one researcher for each clinical outcome (see PICOTS) by following the principles for

adapting GRADE (Grading of Recommendations Assessment, Development and Evaluation) outlined in the AHRQ medical testing methods guide.[18] In determining the strength of a body of evidence regarding a given outcome, the following domains are considered:

- Risk of bias: the extent to which studies reporting on a particular outcome are likely to be protected from bias; graded as low, medium, or high risk of bias)
- Consistency: the extent to which studies report the same direction of effect for a particular outcome; graded as consistent, inconsistent, or unknown (in the case of a single study)
- Directness: reflects whether the outcome is directly or indirectly related to health outcomes of interest; graded as direct or indirect.
- Precision: describes the level of certainty of the estimate of effect for a particular outcome with a precise estimate being on that allows a clinically useful conclusion; graded as precise or imprecise
- Publication bias: indicates that studies may have been published selectively based on consideration of the extent to which relevant empirical findings (e.g., negative or no-difference findings) have not been published or are not available. This is difficult to assess for reviews of diagnostic testing and statistical methods of assessing this may be misleading.[18, 21] Clinical trial registries will be searched for unpublished studies and information from SIPs will be evaluated; graded as suspected or undetected.

Briefly, bodies of evidence consisting of RCTs are initially considered as high strength while bodies of comparative observational studies begin as low strength evidence. The strength of the evidence may be downgraded based on the limitations described above. There are also situations where the observational evidence may be upgraded (e.g., large magnitude of effect, presence of dose-response relationship or existence of plausible unmeasured confounders) as described in the AHRQ methods guides.[15, 18]

A final strength of evidence grade will be assigned by evaluating and weighing the combined results of the above domains. To ensure consistency and validity of the evaluation, the grades will be reviewed by the entire team of investigators. The strength of evidence will be assigned an overall grade of high, moderate, low, or insufficient according to a four-level scale

- High—We are very confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has few or no deficiencies. We believe that the findings are stable, i.e., another study would not change the conclusions.
- Moderate—We are moderately confident that the estimate of effect lies close to the true effect for this outcome. The body of evidence has some deficiencies. We believe that the findings are likely to be stable, but some doubt remains.
- Low—We have limited confidence that the estimate of effect lies close to the true effect for this outcome. The body of evidence has major or

numerous deficiencies (or both). We believe that additional evidence is needed before concluding either that the findings are stable or that the estimate of effect is close to the true effect.

- Insufficient—We have no evidence, we are unable to estimate an effect, or we have no confidence in the estimate of effect for this outcome. No evidence is available or the body of evidence has unacceptable deficiencies, precluding reaching a conclusion.

G. Assessing Applicability

Applicability will be estimated by examining the characteristics of the patient populations (e.g., demographic characteristics such as age, sex, comorbidities); the sample size of the studies; and clinical settings (e.g., academic setting, provider experience) in which the studies are performed. Variability in the studies may limit the ability to generalize the results to other populations and settings.

V. References

1. Ruigomez A, Rodriguez LA, Wallander MA, et al. Chest pain in general practice: incidence, comorbidity and mortality. Fam Pract. 2006;23(2):167-74.
2. Go AS, Mozaffarian D, Roger VL, et al. Heart disease and stroke statistics--2013 update: a report from the American Heart Association. Circulation. 2013;127(1):e6-e245.
3. Lozano R, Naghavi M, Foreman K, et al. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012;380(9859):2095-128.
4. Heidenreich PA, Trogdon JG, Khavjou OA, et al. Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association. Circulation. 2011;123(8):933-44.
5. Detry JM. The pathophysiology of myocardial ischaemia. Eur Heart J. 1996 Dec;17 Suppl G:48-52. PMID: 8960456.
6. Schoen F. Chapter 13, The Heart. In: Cotran RS KV, Collins T, ed Robbins Pathophysiologic Basis of Disease. sixth ed. Philadelphia, London, Toronto, Montreal, Sydney, Tokyo: W. B. Saunders Company; 1999:543-99.
7. Lloyd-Jones D, Adams RJ, Brown TM, et al. Heart disease and stroke statistics--2010 update: a report from the American Heart Association. Circulation. 2010 Feb 23;121(7):e46-e215. PMID: 20019324.
8. Fihn SD, Gardin JM, Abrams J, et al. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, and the American College of Physicians, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J Am Coll Cardiol. 2012;60(24):e44-e164.

9. Fuster V. 50th anniversary historical article. Acute coronary syndromes: the degree and morphology of coronary stenoses. J Am Coll Cardiol. 1999;34(7):1854-6.
10. Topol EJ, Nissen SE. Our preoccupation with coronary luminology. The dissociation between clinical and angiographic findings in ischemic heart disease. Circulation. 1995;92(8):2333-42.
11. Zir LM, Miller SW, Dinsmore RE, et al. Interobserver variability in coronary angiography. Circulation. 1976;53(4):627-32.
12. Tavakol M, Ashraf S, Brener SJ. Risks and complications of coronary angiography: a comprehensive review. Global journal of health science. 2012 Jan;4(1):65-93. PMID: 22980117.
13. Patel MR, Dehmer GJ, Hirshfeld JW, et al. ACCF/SCAI/STS/AATS/AHA/ASNC/HFSA/SCCT 2012 Appropriate use criteria for coronary revascularization focused update: a report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, Society for Cardiovascular Angiography and Interventions, Society of Thoracic Surgeons, American Association for Thoracic Surgery, American Heart Association, American Society of Nuclear Cardiology, and the Society of Cardiovascular Computed Tomography. J Am Coll Cardiol. 2012 Feb 28;59(9):857-81. PMID: 22296741.
14. Shaw LJ, Berman DS, Picard MH, et al. Comparative definitions for moderate-severe ischemia in stress nuclear, echocardiography, and magnetic resonance imaging. JACC. Cardiovascular imaging. 2014 Jun;7(6):593-604. PMID: 24925328.
15. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. AHRQ Publication No. 10(14)-EHC063-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2014. Chapters available at: www.effectivehealthcare.ahrq.gov.
16. Shea BJ, Hamel C, Wells GA, et al. AMSTAR is a reliable and valid measurement tool to assess the methodological quality of systematic reviews. Journal of clinical epidemiology. 2009 Oct;62(10):1013-20. PMID: 19230606.
17. White CM, Ip S, McPheeters M, et al. AHRQ Methods for Effective Health Care Using Existing Systematic Reviews To Replace De Novo Processes in Conducting Comparative Effectiveness Reviews. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008.
18. Methods Guide for Medical Test Reviews. AHRQ Publication No. 12-EC017. Rockville, MD: Agency for Healthcare Research and Quality; June 2012. www.effectivehealthcare.ahrq.gov/reports/final.cfm. Also published as a special supplement to the Journal of General Internal Medicine, July 2012.
19. Higgins JPT GS, ed Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. ed; 2011.
20. Viswanathan M, Ansari MT, Berkman ND, et al. AHRQ Methods for Effective Health Care Assessing the Risk of Bias of Individual Studies in Systematic Reviews of Health Care Interventions. Methods Guide for Effectiveness and Comparative Effectiveness Reviews. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008.
21. van Enst WA, Ochodo E, Scholten RJ, et al. Investigation of publication bias in meta-analyses of diagnostic test accuracy: a meta-epidemiological study. BMC medical research methodology. 2014;14:70. PMID: 24884381.
22. Amsterdam EA, Wenger NK, Brindis RG, et al. 2014 AHA/ACC Guideline for the Management of Patients With Non-ST-Elevation Acute Coronary Syndromes: Executive

Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2014 Sep 18PMID: 25260717.

VI. Definition of Terms

Acute Coronary Syndrome (ACS)
A condition in which a coronary artery becomes blocked by plaque and blood flow to the heart is diminished. In some cases, the plaque may rupture and create a blood clot; this combination of plaque and blood clot may lead to myocardial infarction (a heart attack). ACS refers to a spectrum of clinical presentations ranging from those for ST-segment elevation myocardial infarction (STEMI) to those found in non-ST segment elevation myocardial infarction (NSTEMI).

Atheromatous plaque
An accumulation of fibrous tissues, lipids, and macrophage cells within the artery walls that narrows the artery and restricts blood flow.

Computed tomography (CT) with fractional flow reserve (FFR)
In order to estimate coronary blood flow, mathematical models are applied to data from a CT scan. This has been developed as an alternative to invasive FFR measurement in which a pressure wire is placed in the artery.

Computed tomography (CT) perfusion
Dye is injected into a person's vein and then an X-ray and computer are used to create 3D (three-dimensional) pictures of the heart and arteries to evaluate blood flow and/or damage to the muscle.

Coronary calcium scoring via EBCT (electron beam CT) or MDCT (multidetector CT)
Utilizes X-rays, without the use of intravenous contrast dye, to look for calcium deposits (calcifications) within the coronary arteries that may be obstructing the pathway; the degree and extent of calcification is expressed as a calcium score.

Coronary CT angiography (CCTA)
Utilizes X-rays and contrast dye to produce 3D images of the heart and arteries. CCTA is able to image blockages and narrowing of the arteries without the use of a catheter.

Coronary magnetic resonance angiography (MRA)
Utilizes magnetic fields and pulses or radio wave energy rather than X-rays to view blood vessels. Calcifications, plaque buildup, blood vessel narrowing, and tearing of a vessel are all visible with this imaging technique. May be performed with or without a contrast dye.

Exercise electrocardiogram (ECG) without imaging
A person walks on a treadmill or pedals a stationary bike in order to elevate their heart rate while an ECG records the electrical activity of the heart. Blood pressure and breathing are monitored as well.

Exercise/pharmacologic echocardiography (with or without myocardial echo contrast)
A person's heart is stressed either through exercise or the use of drugs that increase the heart rate and an ultrasound is used to create an image of the heart that is used to evaluate cardiac function and identify any structural abnormalities of the heart.

Guideline-Directed Medical Therapy (GDMT)
Term used to represent medical therapy that is strongly recommended by (primarily Class I and IIa) ACC/AHA guidelines.[8]

Non-ST-Elevation Acute Coronary Syndromes (NSTE-ACS)
This term has been adopted by the American College of Cardiology and the American Heart Association as it "emphasizes the continuum between unstable angina and NSTEMI".[22]

Non-ST-segment elevation myocardial infarction (NSTEMI)
Also called a heart attack and is caused by a partial or temporary blockage, resulting in relatively minimal damage to the heart muscle; it does not cause changes on an ECG but can be demonstrated by an elevation of cardiac biomarkers in the blood.

Pharmacologic stress magnetic resonance imaging (MRI)
Drugs are used to stress the heart and then an MRI, which uses magnets, radiofrequencies and a computer, produces images of the heart and arteries. Contrast may also be used to map the flow of blood.

Positron emission tomography (PET)
A radioactive component is injected into the patient and the uptake and decay of the compound are used to create images of the coronary arteries and the heart and provide information on blood flow. PET and SPECT are similar, but PET provides a higher level of resolution than SPECT.

Predictive accuracy studies
Studies of predictive accuracy use test result information to identify people who will have a *future* event, such as myocardial infarction. Such studies identify patients who benefit from treatment and those do not; may also be referred to as prognostic accuracy.

Serum cardiac biomarkers
Blood serum is made through the process of coagulation to remove blood cells and clotting proteins in order to isolate other proteins like kinases and troponins. This protein-containing serum may then be used to evaluate heart function and aid in diagnosis of a cardiac disorder.

Single photon emission computed tomography (SPECT)
A radioactive component is injected into a person and a gamma camera records emissions of the radiation in a series of 2D images from different angles that are combined to create a 3D view.

ST-segment elevation myocardial infarction (STEMI)

Also called a heart attack and caused by a complete blockage of the coronary artery resulting in a prolonged period of ischemia; it affects a large area of the heart muscle and thus causes changes on an ECG as well as in the serum cardiac biomarkers.

T-wave (ECG change)
When the heart's ventricles become repolarized, this appears on an ECG as the T-wave. Inversion or change in amplitude or symmetry of the T-wave may be an indication of cardiac dysfunction.

VII. Summary of Protocol Amendments

If we need to amend this protocol, we will give the date of each amendment, describe the change and give the rationale in this section. Changes will not be incorporated into the protocol.

Date	Amendment	Rationale
12/18/2014	Exclusion of CT Perfusion. Revisions were made to the Background, the PICOTS, and the Methods sections to reflect this change.	Based on review by internal experts and consultation with the TEP, CT perfusion was excluded as it does not fit within the focus of the review on widely available, established/accepted tests which may commonly be used as a *first* test (aside from resting ECG) for diagnosis of CAD in the population of interest. While it is available from the perspective that it requires no additional hardware or software to do it in a rudimentary fashion, it is not well validated, requires an extra scan (with a vasodilator), more contrast and radiation, it is not currently used outside of the research setting.
12/18/2014	Remove reference to coronary magnetic resonance angiography (MRA) from the Background.	To avoid confusion for readers, reference to MRA in the background was removed. MRA is not an included test for this review as it is an experimental technique and not well established for the diagnosis of CAD.
12/18/2014	Add the following sentence under Clinical Outcomes (primary focus) in the PICOTS section: "For studies of predictive accuracy that do not compare two tests, only the following hard clinical outcomes will be evaluated: MI, heart failure, death; other outcomes listed above	Enhanced clarification regarding inclusion of predictive accuracy studies that do not compare two tests but may compare those who test positive and are treated with those who test negative and are not treated. For these studies, only the following hard clinical outcomes will be evaluated: MI, heart failure,

	will be evaluated based on studies comparing two or more tests"	death. Comparative studies (i.e., those comparing two tests) will be used to evaluate other primary outcomes.
12/18/2014	Add the following sentence under Intermediate Outcomes in the PICOTS section: "to be evaluated based on comparative studies only"	Enhanced clarification. Only comparative studies (i.e., those comparing two tests) will be used to evaluate intermediate outcomes.

VIII. Review of Key Questions

Key questions were reviewed and refined as needed by the Evidence-based Practice Center (EPC) with input from the AHRQ Task Order Officer (TOO) to ensure that the questions are specific and explicit about what information is being reviewed. AHRQ posted the provisional key questions on the Effective Health Care Website for public comment. The EPC refined the key questions after review of the public comments, and input from Key Informants. This input is intended to ensure that the key questions are specific and relevant.

IX. Key Informants

Key Informants are the end users of research, including patients and caregivers, practicing clinicians, relevant professional and consumer organizations, purchasers of health care, and others with experience in making health care decisions. Within the EPC program, the Key Informant role is to provide input into identifying the Key Questions for research that will inform healthcare decisions. The EPC solicits input from Key Informants when developing questions for systematic review or when identifying high priority research gaps and needed new research. Key Informants are not involved in analyzing the evidence or writing the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Key Informants must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their role as end-users, individuals are invited to serve as Key Informants and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

X. Technical Expert Panel

A multi-disciplinary group of clinical, content, and methodologic experts will provide input in further defining populations, interventions, comparisons, or outcomes as well as identifying particular studies or databases to search. They will provide broad expertise and perspectives specific to the topic under development. Divergent and conflicted opinions are common and perceived as health scientific discourse that results in a thoughtful, relevant systematic review. Therefore study questions, design and/or methodological approaches do not necessarily represent the views of individual technical and content experts. Technical Experts will provide information to the EPC to identify literature search strategies and recommend approaches to specific issues as

requested by the EPC. Technical experts do not do analysis of any kind nor contribute to the writing of the report and have not reviewed the report, except as given the opportunity to do so through the peer or public review mechanism.

Technical Experts must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Because of their unique clinical or content expertise, individuals are invited to serve as Technical Experts and those who present with potential conflicts may be retained. The TOO and the EPC work to balance, manage, or mitigate any potential conflicts of interest identified.

XI. Peer Reviewers

Peer reviewers are invited to provide written comments on the draft report based on their clinical, content, or methodologic expertise. Peer review comments on the preliminary draft of the report are considered by the EPC in preparation of the final draft of the report. Peer reviewers do not participate in writing or editing of the final report or other products. The synthesis of the scientific literature presented in the final report does not necessarily represent the views of individual reviewers. The dispositions of the peer review comments are documented and will, for Comparative Effectiveness Reviews and Technical briefs, be published three months after the publication of the Evidence report.

Potential Reviewers must disclose any financial conflicts of interest greater than $10,000 and any other relevant business or professional conflicts of interest. Invited Peer Reviewers may not have any financial conflict of interest greater than $10,000. Peer reviewers who disclose potential business or professional conflicts of interest may submit comments on draft reports through the public comment mechanism.

XII. EPC Team Disclosures

EPC core team members must disclose any financial conflicts of interest greater than $1,000 and any other relevant business or professional conflicts of interest. Related financial conflicts of interest which cumulatively total greater than $1,000 will usually disqualify EPC core team investigators.

XIII. Role of the Funder

This project was funded under Contract No. HHSA 290-2012-00014-I from the Agency for Healthcare Research and Quality, U.S. Department of Health and Human Services. The Task Order Officer reviewed contract deliverables for adherence to contract requirements and quality. The authors of this report are responsible for its content. Statements in the report should not be construed as endorsement by the Agency for Healthcare Research and Quality or the U.S. Department of Health and Human Services.