Comparative Effectiveness Review
Number 171

Effective Health Care Program

Noninvasive Testing for Coronary Artery Disease

Executive Summary

Background

Nature and Burden of Coronary Artery Disease

The public health and economic burdens of coronary artery disease (CAD) are substantial. CAD causes one in six deaths in the United States and is the leading cause of death globally.[1] Annually, approximately 635,000 Americans experience a new coronary event, 280,000 have a recurrent ischemic event, and an additional 150,000 have a silent first myocardial infarction (MI).[2] A large proportion of ambulatory health care visits are for evaluation of patients with suspected CAD, with an estimated 1.5 percent of the population presenting to health care providers with chest pain every year.[3] An estimated $108.9 billion are spent annually on CAD treatment.[4] Optimizing the process for assessing these patients presents an opportunity to improve patient outcomes and target health resources to where they can have the most impact.

The most common underlying cause of CAD is atherosclerosis, a disease process in which plaque builds up on artery walls and can lead to the partial or complete blockage of coronary arteries. As a result, the heart cannot receive adequate blood, oxygen, and vital nutrients. Plaque causes blockage by two mechanisms: (1) progressive narrowing of the artery because plaque compromises the vessel lumen and (2) thrombotic occlusion

Effective Health Care Program

The Effective Health Care Program was initiated in 2005 to provide valid evidence about the comparative effectiveness of different medical interventions. The object is to help consumers, health care providers, and others in making informed choices among treatment alternatives. Through its Comparative Effectiveness Reviews, the program supports systematic appraisals of existing scientific evidence regarding treatments for high-priority health conditions. It also promotes and generates new scientific evidence by identifying gaps in existing scientific evidence and supporting new research. The program puts special emphasis on translating findings into a variety of useful formats for different stakeholders, including consumers.

The full report and this summary are available at **www.effectivehealthcare.ahrq.gov/reports/final.cfm**.

AHRQ
Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Effective Health Care

of the artery, which occurs when the hard surface of a plaque tears or breaks off and exposes the inner fatty prothrombotic and platelet-attracting components to the site, resulting in enlargement of the blockage. The resulting reduction in blood flow can be either acute or chronic and leads to an imbalance in the blood supply to the myocardium, thus increasing the requirements of the myocardium for oxygenated blood either at rest or during exertion.[5,6]

The most common symptom of obstructive CAD is chest pain (angina), which is the first presenting symptom in at least 50 percent of patients with CAD.[7] Other common symptoms include the angina equivalents dyspnea, early fatigue with exertion, indigestion, palpitations, tightness in the throat, and neck or arm pain. However, because these symptoms are also seen in many common noncardiac conditions, such as gastroesophageal reflux, esophageal spasm, and cervical disc disease, they are much less reliable predictors of CAD. Women and people with diabetes are less likely to experience classic angina, making early diagnosis of CAD challenging in these populations. The onset of symptoms and clinical impact of CAD depend on a variety of factors, including plaque distribution and degree of vessel narrowing; however, lesion severity does not necessarily correlate well with symptoms. Further, CAD may remain asymptomatic for many years.

Diagnosis of CAD

Accurate early diagnosis of CAD in symptomatic patients is important for initiation of appropriate treatment and reduction of CAD-related morbidity and mortality. Diagnosis of CAD begins with a thorough clinical workup, including a physical examination, patient history, and possibly resting electrocardiography (ECG), followed by noninvasive testing if in an outpatient clinic. In addition to physical examination and patient history for people presenting with chest pain to the emergency department (ED), some combination of a resting ECG, chest x ray, and/or serum biomarkers such as cardiac troponins is generally done. If the presentation is not acute, the ECG is nonspecific, and cardiac troponins are normal, then the stable patient may be discharged or receive further testing to help determine the etiology of chest pain and the appropriate management. Patients with a high suspicion for a noncardiac etiology of chest pain may forgo evaluation for occlusive CAD or ischemia in favor of pursuing other testing for such causes (e.g., pulmonary embolism).

A diagnosis of CAD can be made by looking for evidence of the pathophysiologic processes of disease, including anatomic changes of the arterial wall, impaired myocardial perfusion, or consequences of impaired perfusion, such as myocardial contractile dysfunction. Historically, invasive coronary angiography (ICA) has been considered the standard reference diagnostic test for anatomic CAD, defined here as any obstructive lesion that is consistent with symptoms or that may carry an increased risk of acute coronary syndrome (ACS), although its invasive nature makes it less ideal in many patients because of its associated risks and costs. Noninvasive tests are another option, and provide diagnostic and prognostic information that can improve risk stratification, thus guiding subsequent testing and interventions. Noninvasive diagnostic tests can be broadly divided into two categories: functional tests and anatomic tests. Functional tests provide information not provided by standard ICA, such as whether symptoms are correlated with areas of ischemia. Functional tests include exercise ECG, exercise/pharmacologic stress echocardiography, exercise/pharmacologic cardiac nuclear imaging with single-photon emission computed tomography (SPECT) or positron emission tomography (PET), pharmacologic stress magnetic resonance imaging (MRI), computed tomography (CT), and Doppler ultrasound–derived flow reserve measurements. Noninvasive anatomic tests include coronary CT angiography (CCTA) and coronary artery calcium scoring (CACS). American College of Cardiology Foundation/American Heart Association (ACCF/AHA) Appropriate Use Criteria suggest that, as a general rule, functional testing is more informative than noninvasive anatomic evaluation and exercise testing is more informative than pharmacologic testing.[8]

Deciding which test to use for diagnosis of CAD in stable symptomatic patients is not a simple matter. A patient's pretest CAD risk can be informative as to the test or procedure most appropriate as a first step toward diagnosing CAD. While there are a number of standard risk-assessment tools, these are rarely documented in clinical practice, and the clinician's overall assessment of sociodemographic characteristics (e.g., sex, age) and characteristics of the chest pain (typical or atypical) is the most common assessment of pretest likelihood of CAD. Pretest risk of CAD is frequently based on the ACCF/AHA Guideline and defined as low (<10% pretest probability of CAD), intermediate (10%–90% pretest probability of CAD), or high (>90% pretest probability of CAD).[9] Patients at low pretest risk may undergo noninvasive testing to further delineate their risk and to provide a basis for clinical decisionmaking, although in some cases, an alternative explanation for the symptoms (such as

heartburn, costochondritis, or pulmonary disease) may be evaluated first. Patients at intermediate risk commonly undergo noninvasive testing, followed by appropriate treatment for comorbidities and risk factors. The ACCF/AHA intermediate range is intentionally broad, reflecting the availability of noninvasive tests that have been viewed as both safe and effective to further stratify risk in the intermediate pretest risk category. In other words, the low end of the intermediate range is extended irrespective of cost because of the important health consequences of missing disease, but this also results in a situation in which testing is performed in a very large number of individuals who do not have disease.[10] The high end is extended because of the combination of the somewhat high cost and risk of ICA and reasonably high sensitivity of testing to detect high-risk obstructive disease. Patients at high risk may undergo noninvasive testing, although at times clinicians may appropriately decide to bypass noninvasive stress testing and proceed directly to ICA.[8] This is more frequently done in patients who present to the ED with typical symptoms. In patients for whom clinical judgment remains equivocal, an additional test to further identify risk may be pursued.

The 2012 ACCF/AHA Guideline states that diagnostic testing is most valuable when the pretest probability of ischemic heart diseases is intermediate (10%–90%) and provides a range of options for tests that may be used in a given scenario. However, the effectiveness of different modalities with regard to impact on clinical outcomes is not compared.[9] There remains uncertainty regarding which tests, if any, may be most suitable and most beneficial for specific scenarios in patients who present with symptoms suggestive of CAD. Specifically—

- In patients with low pretest probability of CAD (<10%), are clinical outcomes improved by use of stress testing with or without imaging or with no further testing? It is not clear whether imaging may be necessary in this group of patients, or if there are specific subgroups of low-risk patients who might benefit more from one type of testing than another or who should have no further testing.
- How do tests compare with regard to improvement in clinical outcomes (e.g., MI, premature mortality, and congestive heart failure) in patients whose risk is very low (<5%) or in patients with intermediate to high risk? How do tests differ in their ability to reclassify patient risk after the test and to influence appropriate patient management?
- Are there differences in clinical outcomes following anatomic versus functional testing in either the low-risk group or the group with intermediate to high risk?

Scope and Key Questions

The objective of this review is to assess the effectiveness of noninvasive technologies for the diagnosis of CAD or dysfunction that results in symptoms attributable to myocardial ischemia in patients who present with signs or symptoms suggestive of CAD, whose condition is considered to be stable, and who have no known history of CAD. The intended focus is on clinical outcomes and clinical pathways following the first diagnostic test performed as a result of initial risk assessment, which includes clinical presentation and physical exam, family history of CAD, and findings on resting ECG. Further, this report focuses on established tests for diagnosing CAD. Harms related to both the initial test and subsequent testing are evaluated. Information on the traditional measures of accuracy (e.g., sensitivity and specificity) of noninvasive tests versus the historically accepted gold standard of ICA comprises the majority of the literature and is presented for context. Increasingly, experts in cardiovascular health indicate that evidence on the value of noninvasive diagnostic cardiovascular testing needs to expand beyond traditional measures of test performance, such as sensitivity and specificity compared with a given reference standard, and focus on evaluating the impact of such testing on hard cardiovascular outcomes and downstream harms. Thus, while diagnostic accuracy measures provide important information on test performance, the primary focus of this report is to determine whether noninvasive tests improve clinical health outcomes and impact patient management.

The analytic framework (Figure A) shows the target population, interventions, and outcomes that were examined.

The Key Questions for this Comparative Effectiveness Review are as follows.

In stable symptomatic patients with suspected CAD who do not have previously diagnosed CAD and who have had a resting ECG—

- For patients considered to be at ***very low or low risk*** for CAD, what is the comparative effectiveness of **anatomic tests (compared with each other, usual care, or no testing)?**

- For patients considered to be at ***very low or low risk*** for CAD, what is the comparative effectiveness of **functional tests (compared with each other, usual care, or no testing)?**
- For patients considered to be at ***intermediate to high risk*** for CAD, what is the comparative effectiveness of **anatomic tests (compared with each other, usual care, or no testing)?**
- For patients considered to be at ***intermediate to high risk*** for CAD, what is the comparative effectiveness of **functional tests (compared with each other, usual care, or no testing)?**
- What is the comparative effectiveness of **anatomic tests versus functional tests** in those who are at ***very low or low risk*** for CAD?
- What is the comparative effectiveness of **anatomic tests versus functional tests** in those who are at ***intermediate to high risk*** for CAD?

For each Key Question, the following subquestions were explored:

a. What is the effectiveness of the compared tests for improving primary clinical health outcomes (e.g., quality of life, avoiding MI)?

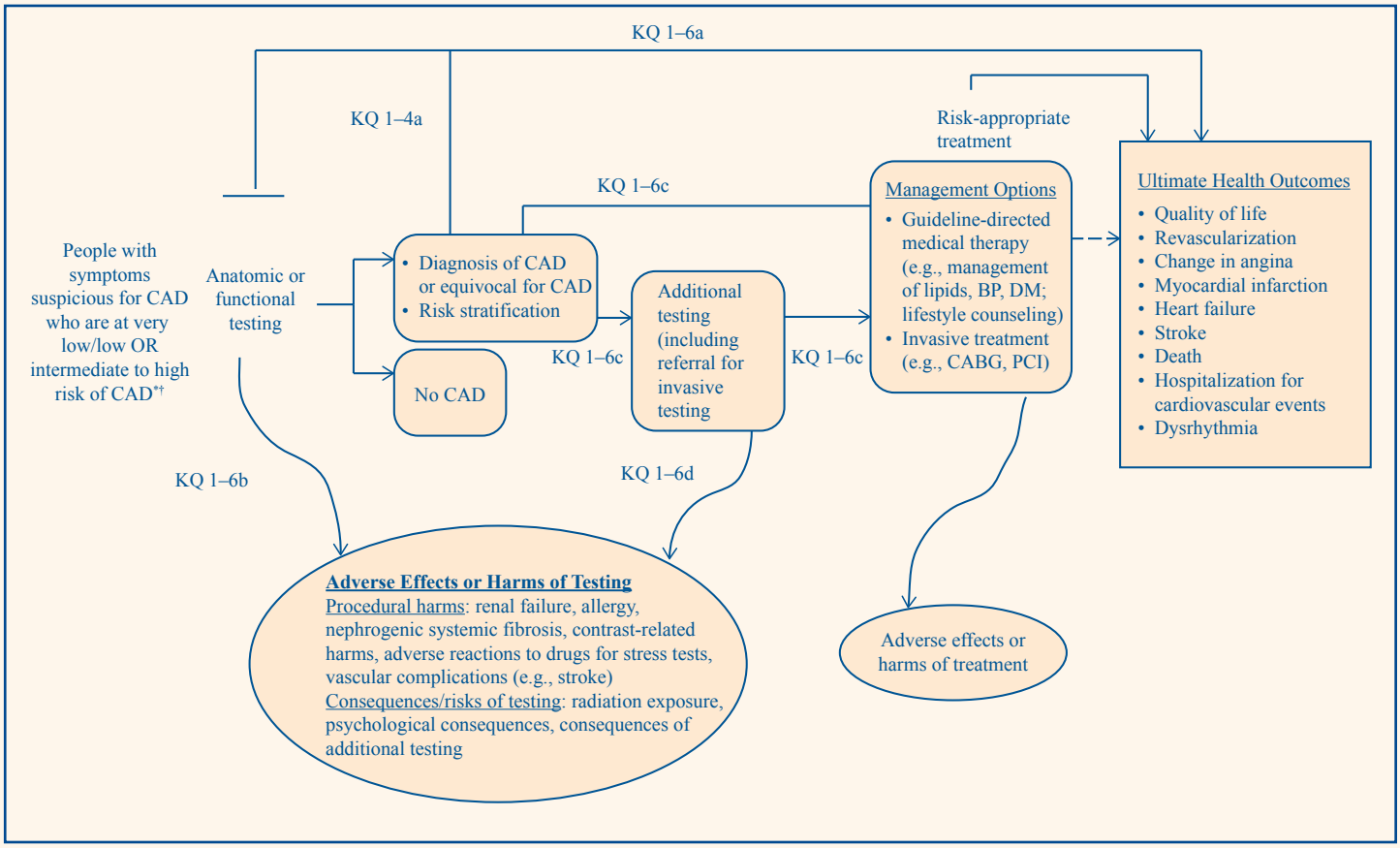
b. What are the adverse effects, consequences, or harms of testing?

c. How do noninvasive tests differ in terms of clinical management based on test results, including referral for coronary angiography or additional noninvasive testing?

d. What harms are associated with additional testing following anatomic tests?

e. Is there differential effectiveness or harm based on patient characteristics (e.g., sex, age, comorbidities) or the patient's ability to exercise?

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



BP = blood pressure; CABG = coronary artery bypass graft; CAD = coronary artery disease; DM = diabetes mellitus; KQ = Key Question; PCI = percutaneous coronary intervention

*People at very low or low risk are evaluated separately from those at intermediate to high risk when possible.

†KQ 1–6e: Potential modifiers related to differential efficacy and/or safety include patient factors (e.g., age, sex), comorbidities, and ability to exercise.

Methods

The methods for this Comparative Effectiveness Review follow the guidance in the Agency for Healthcare Research and Quality (AHRQ) "Methods Guide for Effectiveness and Comparative Effectiveness Reviews" (Methods Guide).[11]

Topic Refinement and Review Protocol

The topic for this Comparative Effectiveness Review was ranked as a priority topic by a panel of stakeholders convened through the Duke Evidence-based Practice Center's Cardiovascular Topic Identification project. The preliminary Key Questions were posted on AHRQ's Web site for public comment for 4 weeks. Public comments and input from the Technical Expert Panel (TEP) were used to develop the final Key Questions and protocol. The TEP, convened to provide high-level content and methodological guidance to the review process, consisted of experts in cardiology and cardiac diagnostic testing, radiology, internal medicine, and health services research, as well as professional organizations and policymakers. TEP members disclosed all financial or other conflicts of interest prior to participation. The AHRQ Task Order Officer and the investigators reviewed the disclosures and determined that the TEP members had no conflicts of interest that precluded participation.

Both the final topic-refinement document and the systematic review protocol, developed prior to initiation of the review, can be found on the AHRQ Web site at www.effectivehealthcare.ahrq.gov/search-for-guides-reviews-and-reports/. The protocol is also registered with the PROSPERO international database of prospectively registered systematic reviews (CRD42015022081).

Literature Search Strategy

A research librarian conducted searches for primary studies in the following databases through July 2015: Ovid MEDLINE®, Cochrane Central Register of Controlled Trials, Cochrane Database of Systematic Reviews, and Evidence-Based Medicine Reviews–Health Technology Assessment. A search strategy was developed based on an analysis of the medical subject heading (MeSH) terms and text words of key articles identified a priori. (The full search strategy is available in Appendix A of the full report.) Search start dates were not restricted. The reference lists of included articles and relevant review articles were also reviewed. All citations were downloaded and imported into an electronic database (EndNote® X7, Thomson Reuters, Philadelphia, PA). A list of relevant drugs and manufacturers was provided to the Scientific Resource Center, which requested Scientific Information Packets, and relevant published and unpublished studies were assessed for inclusion in the final report. Additional details regarding handling of citations are found in the full report and in Appendix A of the full report.

Literature searches were updated during the public comment and peer review period in order to ensure that any new publications that met our inclusion criteria were incorporated into the final report.

Inclusion and Exclusion Criteria

Criteria for inclusion and exclusion of studies were based on the Key Questions and the PICOTS (populations, interventions, comparators, outcomes, timing, and setting) approach. Studies of stable symptomatic adult patients undergoing their first noninvasive diagnostic test for suspected CAD were sought. Studies of patients with known CAD (prior MI or prior revascularization) were excluded. In keeping with the review protocol, studies of patients with definite ACS, non–ST-elevation acute coronary syndromes, non–ST-segment elevation myocardial infarction (NSTEMI), and ST-segment elevation myocardial infarction (STEMI) were excluded (or were included only if these patients did not comprise >20% of the study population), as were studies of patients with unstable angina and elevated serum cardiac biomarkers or ECG changes. For all Key Questions, the focus was on evidence from comparative studies with the least potential for bias. Noncomparative studies of predictive accuracy were considered if there was a lack of comparative data for a specific diagnostic modality. Interventions of interest included anatomic imaging (i.e., CCTA, coronary calcium scoring via electron beam or multidetector CT) and functional tests (i.e., stress ECG, stress echocardiography, stress nuclear imaging [SPECT, PET], and stress MRI). Comparators included other noninvasive tests included in the interventions, usual care (as defined by the authors), or no testing. Studies that included technologies that are not widely available, are no longer used, or have not been established for the diagnosis of CAD were excluded.

The primary outcomes (see "Rating the Body of Evidence" section) were considered to be the most clinically important and were the focus of reporting, decisions for data pooling, and determination of overall strength of evidence. Additional outcomes are reported in the detailed evidence synthesis sections of the Results chapter of the full report, organized by the Key Questions, with a focus on outcomes common across studies. Where applicable and where data were available, results from the index visit

and the followup period were reported separately. For studies of predictive accuracy, only hard clinical outcomes (i.e., MI, death, composite cardiac outcome, heart failure) were evaluated. For both the initial test and any subsequent downstream testing, the primary safety outcomes were related to harms of testing (e.g., adverse reaction or allergy to contrast or stress agents) and risks and consequences of testing (e.g., radiation exposure). Studies focused on "per-vessel" or "per-segment" analysis without per-patient findings were excluded, and treatments and outcomes of treatments were beyond the scope of this report. Studies published only as conference abstracts, non–English-language articles, and studies of nonhuman subjects were excluded. Studies had to report original data to be included.

Study Selection

Abstracts for all citations from the literature searches were independently reviewed by two team members and results were recorded in EndNote. All citations that either reviewer found to be potentially appropriate for inclusion underwent full-text review. Two investigators independently evaluated each full-text article for final inclusion. For inclusion, both reviewers had to agree that inclusion criteria were met. Differences between reviewers were resolved through consensus and discussion. A record of studies excluded at the full-text level with reasons for exclusion is included in Appendix C of the full report.

Data Extraction

The investigative team created a form in Microsoft® Excel for abstracting the data elements for the Key Questions. Two staff members and five experienced team members entered data. After data extraction, at least one other staff member and one investigator verified the accuracy and completeness of abstraction. Discrepancies were resolved by discussion and consensus. Specific information included in the data extraction forms is outlined in Appendix D of the full report.

Quality (Risk-of-Bias) Assessment of Individual Studies

Predefined criteria were used to assess the quality (risk of bias) of included randomized controlled trials (RCTs) and observational studies by using clearly defined templates and criteria as appropriate and following guidance from the AHRQ Methods Guide.[11] Assessment of RCTs followed appropriate criteria and methods established in the *Cochrane Handbook for Systematic Reviews of Interventions*.[12] Comparative observational studies were assessed for study design features and sources of potential bias. These criteria and methods were used in concordance with the AHRQ schema, and each study was rated as being "good," "fair," or "poor" quality.[13]

Studies rated "good" are considered to have the least risk of bias, and their results are considered valid. Studies rated "fair" are susceptible to some bias, although not enough to invalidate the results. The fair-quality category is broad, and studies with this rating vary in their strengths and weaknesses. Studies rated "poor" have significant flaws that imply biases of various types that may invalidate the results. 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 were present.

Each study evaluated was independently reviewed for quality by two team members. Any disagreements were resolved by consensus. The final quality assessments are described in detail in Appendix I of the full report.

Data Synthesis

When adequate data were reported in at least two studies, meta-analysis was conducted in order to provide more precise estimates for outcomes. To determine the appropriateness of conducting meta-analysis, clinical and methodological diversity and assessed statistical heterogeneity were considered. Given the multiple interventions included in this report, a network meta-analysis was planned to estimate the relative effects of interventions that were not directly compared, and to make full use of both direct and indirect evidence.[14] However, the number of included studies turned out to be very small (2 for each comparison), with a limited number of comparisons (only CCTA vs. SPECT and CCTA vs. usual care). Along with heterogeneity across studies, this made network meta-analysis impossible. Therefore, only standard meta-analysis was conducted and only binary outcomes were eligible. The profile-likelihood random-effects model[15] was used to combine risk differences while incorporating variation among studies. The presence of statistical heterogeneity among the studies was assessed by using the standard Cochran's chi-square test, and the magnitude of heterogeneity was assessed by using the I^2 statistic.[16]

To account for clinical heterogeneity, analyses were stratified by pretest risk. Within each stratum, the number of studies was too small for exploring heterogeneity based on any study-level characteristics. Sensitivity analyses using risk ratios were conducted to check the robustness of results to the choice of effect measure. Conclusions were

generally similar and not separately reported. All analyses were performed using Stata®/IC 12.1 (StataCorp, College Station, TX).

Rating the Body of Evidence

The following outcomes were considered to be the most relevant and were the focus of reporting, data pooling, and determination of overall strength of evidence: mortality (all cause), MI, additional noninvasive testing, referral for ICA, and subsequent revascularization (i.e., percutaneous coronary intervention [PCI] or coronary artery bypass graft [CABG]). Primary safety outcomes of interest for both the index test and any subsequent downstream testing included harms of testing (e.g., renal failure, allergic reactions, and adverse reactions to contrast or stress agents) and risk and consequences of testing (e.g., radiation exposure, psychological consequences of diagnosis, incidental findings).

The strength of evidence (high, moderate, low, or insufficient) for each primary effectiveness and safety outcome was initially assessed by one researcher.[11,13] To ensure consistency and validity of the evaluation, the strength-of-evidence ratings for all key outcomes were reviewed by multiple investigators, and discrepancies were resolved by consensus. Bodies of evidence consisting of RCTs started as high strength (greatest confidence that the evidence reflects the true effect; further research is unlikely to change our confidence in the effect estimate), while bodies of comparative observational studies began as low-strength evidence (low confidence in the estimate; further research is likely to change the effect estimate and change the confidence in the estimate). The strength of the evidence was then downgraded based on study limitations (i.e., risk of bias, consistency of effect, directness of outcome, precision of effect estimate, and reporting bias).[11] There are also situations in which the observational evidence may be upgraded (e.g., very large size of effect), but we found no instances in which these could be applied in this body of evidence.[11,17] The detailed strength-of-evidence tables and detailed explanations of the various grades can be found in Appendix J of the full report.

Applicability

Applicability of the evidence was considered by examining the characteristics of the patient populations included in studies (e.g., demographic characteristics, presence of relevant cardiac risk factors, and pretest risk for CAD), the sample size of the studies, and the clinical settings in which the studies were performed (e.g., outpatient clinic, ED), as outlined in the AHRQ Methods Guide.[11,18] Variability in the studies may limit the ability to generalize the results to other populations and settings. For example, older studies of established tests may not be as applicable in light of advances in technology, and short-term outcomes based on immediate decisionmaking in the ED may not be generalizable to longer term outcomes and decisionmaking in the outpatient setting.

Peer Review and Public Commentary

Experts in the diagnosis and treatment of CAD, as well as individuals representing other important stakeholder groups, were invited to provide external peer review of this Comparative Effectiveness Review. The AHRQ Task Order Officer and an Evidence-based Practice Center Program Associate Editor also provided comments and editorial review. The draft report was published on the AHRQ Web site for 4 weeks in order to solicit public comments. At the end of this period, the authors considered both the peer and public review comments and generated a final report. A disposition-of-comments report detailing the authors' responses to the peer and public review comments will be made available 3 months after AHRQ posts the final report on the public Web site.

Results

Results of Literature Searches

Database searches identified 17,146 potentially relevant citations. After dual review of abstracts and titles, 310 articles underwent full-text review; of these, 46 studies (in 51 publications) were determined by dual review to meet the inclusion criteria and were included in this report. The evidence base in this report includes data from RCTs as well as observational studies and noncomparative studies. Studies designed to compare one noninvasive test with another, with usual care, or with no testing form the primary basis for our report.

Organization of Results

Given the heterogeneity in how pretest risk was measured and defined across the studies, results could not be reported as delineated by the Key Questions into distinct pretest risk groups (i.e., low risk and intermediate to high risk). Therefore, the results were organized by pretest risk as defined by the study authors, which included populations with low risk, intermediate risk, low to intermediate risk, intermediate to high risk, high risk, and mixed risk (or pretest risk not reported). Studies describing high pretest risk excluded patients with ACS (or if included, those with ACS comprised <20% of the population) and were

interpreted as representing the higher risk end of the intermediate pretest risk range. Available data from studies conducted in EDs were primarily for the index ED visit and are noted. Outcomes such as MI at the time of the ED index visit were considered to reflect diagnosis of MI at that time. Where available, data on longer term followup are presented. An overview of tests compared for the various pretest risk groups is found in Table A.

Evidence for all outcomes in the low and high pretest risk groups was rated as insufficient; this evidence is not summarized here but is presented in the full report. Evidence for other comparators and primary outcomes considered to be insufficient to draw conclusions because of study limitations and/or imprecision in observational studies or lack of evidence are also available in the full report.

Primary results described here and in Tables B–E are organized by tests compared. Additional detailed results are organized by primary outcomes in the full report in Tables 8–15.

Low Pretest Risk of CAD

A total of two RCTs were identified in populations with a low pretest risk of CAD: CCTA versus usual care (1 RCT)[19] and SPECT versus exercise ECG (1 RCT).[20] Evidence was based on subgroup analyses and was insufficient for all outcomes. Details of these studies are found in the full report.

Intermediate Pretest Risk of CAD

A total of seven comparative studies (in 9 publications) were identified in populations with an intermediate pretest risk of CAD: CCTA versus usual care (2 RCTs,[19,21,22] 1 prospective observational study[23,24]), CCTA versus various functional tests (1 RCT),[25] CCTA versus SPECT,[26] and SPECT versus exercise ECG (2 RCTs).[20,27] Table B summarizes the primary findings for this risk category.

CCTA Versus Usual Care

In intermediate-risk patients presenting to the ED, there was low-strength evidence from two fair-quality trials (N = 1,111) that patients in the CCTA and usual-care groups had similar mortality (≤30 days: 0% in both groups); MI (index ED visit: 2.3% vs. 3.6%; 28 days: 0.2% vs. 0.8%); any revascularization (index ED visit: 7.2% vs. 5.6%); PCI (index ED visit: 5% vs. 3%; 28 days: 0.6% in both groups); CABG (index ED visit: 1% in both groups; 28 days: 0% in both groups); and additional testing at the index ED visit and through 28–30 days (28 days: SPECT [1.6% vs. 1.8%], stress echocardiography [0% in both groups], or exercise treadmill testing [2% vs. 3%]). ICA referral was also similar at the index ED visit (13.8% vs. 11.2%; pooled risk difference [RD], 3; 95% confidence interval [CI], 0 to 7 per 100 patients; $I^2 = 0\%$) and after the index visit through 28 days (1.0% vs. 0.8%) (low strength of evidence).

SPECT Versus Exercise ECG

In 824 intermediate-risk women (setting not reported), groups were similar with respect to mortality (1.0% vs. 0.5%), ICA referral (6% in both groups), revascularization (2.0% vs. 1.0%), and hospitalization for chest pain (3.9% vs. 3.1%) through 24 months, based on one fair-quality trial (low strength of evidence). However, moderate-strength evidence from this trial suggests that SPECT is associated with less additional noninvasive testing than exercise ECG (9.4% vs. 18.6%; RD, -9; 95% CI, -14 to -4 per 100 people). Among those randomized to exercise ECG, the frequency of crossover to SPECT (counted as use of an additional test) was 8, 25, and 43 percent for women who had normal, indeterminate, and abnormal ECG results, respectively. Of those randomized to SPECT, this test was repeated in 9, 8, and 15 percent of women with normal, mildly abnormal, and moderately to severely abnormal results, respectively.

A second fair-quality trial reported that in a subgroup of 280 intermediate-risk outpatients, SPECT was associated with fewer referrals to ICA (10.6% vs. 43.1%; RD, -32; 95% CI, -43 to -22 per 100 people) (low strength of evidence) and additional stress testing (0% vs. 38%; RD, -38; 95% CI, -48 to -29 per 100 people) (low strength of evidence) through a mean of 22 months of followup.

Differences in patient characteristics between the two trials may partially explain differences in findings; one trial was comprised of women with a mean age of 63 years who were able to perform ≥5 METs (a measure of energy expenditure) on the Duke Activity Status Index. Findings from the other trial are based on subanalysis of intermediate-risk patients from a general population composed of more than 50 percent men with mean age of 59 years with any activity ability.

CCTA Versus Functional Testing

In a good-quality trial of 10,003 intermediate-risk outpatients (mean, 53% ± 21% combined Diamond and Forrester and Coronary Artery Surgery Study risk score for likelihood of obstructive CAD), moderate-strength evidence suggested that there was no difference between

Table A. Overview of test comparisons described in the full report and pretest risk groups for which the comparisons were made

Comparator	Pretest Risk Groups With Usual Care Comparisons	Pretest Risk Groups With CCTA Comparisons	Pretest Risk Groups With SPECT Comparisons	Pretest Risk Groups With Stress ECG Comparisons	Pretest Risk Groups With Functional Testing Comparisons	Pretest Risk Groups With Stress Echo-cardiography Comparisons	Pretest Risk Groups With Nuclear MPI Comparisons
Usual Care	N/A	• Low • Intermediate • Low to intermediate • Intermediate to high • High (non-ACS) • Mixed population	No comparison	No comparison	No comparison	No comparison	No comparison
CCTA	• Low • Intermediate • Low to intermediate • Intermediate to high • High (non-ACS) • Mixed population	N/A	• Intermediate • Low to intermediate • Intermediate to high • Mixed population	• Low • Intermediate • Low to intermediate • Intermediate to high • Mixed population	• Intermediate	• Mixed population	• Mixed population
SPECT	No comparison	• Intermediate • Low to intermediate • Intermediate to high • Mixed population	N/A	• Intermediate • High (non-ACS) • Mixed population	No comparison	No comparison	No comparison

Table A. Overview of test comparisons described in the full report and pretest risk groups for which the comparisons were made (continued)

Comparator	Pretest Risk Groups With Usual Care Comparisons	Pretest Risk Groups With CCTA Comparisons	Pretest Risk Groups With SPECT Comparisons	Pretest Risk Groups With Stress ECG Comparisons	Pretest Risk Groups With Functional Testing Comparisons	Pretest Risk Groups With Stress Echo-cardiography Comparisons	Pretest Risk Groups With Nuclear MPI Comparisons
Stress ECG	No comparison	• Low • Intermediate • Low to intermediate • Intermediate to high • Mixed population	• Intermediate • High (non-ACS) • Mixed population	N/A	No comparison	• Mixed population	• Mixed population
Functional Testing	No comparison	• Intermediate	No comparison	No comparison	N/A	No comparison	No comparison
Stress Echocardiography	No comparison	• Mixed population	No comparison	• Mixed population	No comparison	N/A	• Mixed population
Nuclear MPI	No comparison	• Mixed population	No comparison	• Mixed population	No comparison	• Mixed population	N/A

ACS = acute coronary syndrome; CCTA = coronary computed tomography angiography; ECG = electrocardiography; MPI = myocardial perfusion imaging; N/A = not applicable; SPECT = single-photon emission computed tomography

groups in all-cause mortality (12 months: 0.42% vs. 0.64%; median 25 months: 1.48% vs. 1.50%); nonfatal MI (12 months: 0.36% vs. 0.54%; median 25 months: 0.60% vs. 0.80%); or cardiac hospitalizations (median 25 months: 1.22% vs. 0.92%). There was high strength of evidence that CCTA was associated with more ICA referrals (12.19% vs. 8.11%; RD, 4.08; 95% CI, 2.90 to 5.26 per 100 people) and revascularizations (6.22% vs. 3.16%; RD, 3.07; 95% CI, 2.24 to 3.90 per 100 people), including CABG and PCI evaluated separately, through 90 days. Major procedural complications were rare and similar between groups—procedural stroke (0.02% vs. 0.04%), major bleeding (0.1% in both groups), anaphylaxis or renal failure requiring dialysis (no cases) (moderate strength of evidence).

CCTA Versus SPECT

In a fair-quality trial of 400 intermediate-risk patients admitted to a telemetry ward (mean Diamond and Forrester pretest risk of 37%; mean Thrombolysis in Myocardial Infarction [TIMI] score of 1.3 ± 1.0), low-strength evidence suggested that there was no difference between CCTA and SPECT groups in all-cause mortality through a median of 24.5 months (0.5% vs. 3.0%; RD, -2.5; 95% CI, -5.1 to 0.06 events per 100 people) or in 12-month ICA referral (15.0% vs. 16.0%), additional testing (22.5% in both groups), revascularization (7.5% vs. 6.0%), or PCI (4.0% vs. 5.5%). However, CABG was more common following CCTA than SPECT through 12 months (3.5% vs. 0.5%; RD, 3.0; 95% CI, 0.3 to 5.7 events per 100 people), and cardiac rehospitalization occurred in fewer CCTA than SPECT patients through a median of 40.4 months, although the difference did not achieve statistical significance (25.0% vs. 31.0%; RD, -5.5; 95% CI, -14.3 to 0.03 events per 100 people) (low strength of evidence). No major complications were attributed to the imaging procedure; 30-day death, MI, and stroke were not reported. The composite of periprocedural chest pain, shortness of breath, or palpitations occurred in significantly fewer CCTA than SPECT patients (0.5% vs. 15.9%; RD, -15.4; 95% CI, -20.8 to -10.1 per 100 people), while there were no differences between groups in minor adverse reactions, including headache, nausea, dizziness, or feeling of warmth (24.2% vs. 24.5%) or in rash or pruritus (1.6% vs. 0%). There were no cases of post-test renal dysfunction (low strength of evidence).

Low to Intermediate Pretest Risk of CAD

A total of eight comparative studies (in 9 publications) were identified in populations with low to intermediate pretest risk of CAD: CCTA versus usual care (2 RCTs,[28,29] 1 retrospective observational study[30]), SPECT (2 RCTs,[31,32] 1 retrospective observational study[33]), and exercise ECG (1 RCT,[34] 1 retrospective observational study[35,36]). Table C summarizes the primary findings for this risk category.

CCTA Versus Usual Care

A fair-quality trial of 1,370 low- to intermediate-risk patients presenting to the ED (TIMI risk score, 0 [51%], 1 [36%], and ≥2 [13%]) showed no difference between CCTA and usual-care groups in mortality through 1 month (0% in both groups) or MI diagnosis at the index ED visit (1.0% vs. 0.9%) and through 1 month (1.1% in both groups) (low strength of evidence). Moderate-strength evidence from the same trial suggested that CCTA patients were less likely to be hospitalized at the index visit (50% vs. 77%; RD, -26.8; 95% CI, -31.9 to -21.8 per 100 people), but cardiac-related hospitalizations through 1 month were similar (3% vs. 2%). The CCTA groups were less likely to undergo additional testing at the index visit (13.7% vs. 57.8%; RD, -44.1; 95% CI, -49.2 to -39.1 per 100 people) and through 1 month (23.1% vs. 66.4%; RD, -43.3; 95% CI, -48.4 to -38.1 per 100 people) in the same trial (moderate strength of evidence), and through 3 months (33% vs. 60%; RD, -27; 95% CI, -51 to -2) in one poor-quality trial of 60 patients with risk scores not reported (low strength of evidence). ICA referrals were similar for the groups at the index ED visit (4.1% vs. 3.9%; 1 trial; N = 1,392) and through 1- to 3-month followup in two trials (N = 1,452; pooled estimate, 5.2% vs. 4.7%; RD, 1; 95% CI, -1 to 3 per 100 people). There were slightly more revascularization procedures in the CCTA group at the index visit in the larger trial (2.5% vs. 0.9%; RD, 1.7; 95% CI, 0.3 to 3.0 per 100 people), but revascularization frequency was similar through the followup period across both trials (pooled estimate, 2.7% vs. 1.2%; RD, 1; 95% CI, 0 to 3 per 100 people) (low strength of evidence).

CCTA Versus Exercise ECG

Based on one fair-quality trial of 562 low- to intermediate-risk ED patients, there was low-strength evidence of no differences in mortality through 12 months (0.6% vs. 0.4%) or in diagnosis of MI at the index ED visit (1.9% vs. 1.7%) and through 1 month (no additional cases). The 12-month rates of referral to ICA (9.0% vs. 2.3%; RD, 4.8; 95% CI, 0.8 to 8.9 per 100 patients) and revascularization (4.3% vs. 1.3%; RD, 3.1; 95% CI, 0.5 to 5.7 per 100 patients) were significantly greater following CCTA than exercise ECG (low strength of evidence).

CCTA Versus SPECT

In low- to intermediate-risk patients presenting to the ED (median TIMI score, 1.0), there was low-strength evidence from two trials (N = 952; 1 good and 1 fair quality) of no difference through 6 months in mortality (0% in both groups). There was moderate-strength evidence that there was no difference in MI (diagnosis at index ED visit: 0.3% vs. 1.5%; RD, -1.2%; 95% CI, -2.6% to 0.19%; 6 months: 0% in both groups), as reported by both RCTs, or in cardiac-related hospitalizations (0% in both groups), as reported in one good-quality RCT. Together, the trials of ED patients reported that ICA referrals were similar at both the index ED test (7.6% vs. 5.5%; pooled RD, 4; 95% CI, -4 to 11 per 100 patients; $I^2 = 71.7\%$) and through 6 months (0.7% vs. 1.3%; pooled RD, -1; 95% CI, -5 to 3 per 100 patients; $I^2 = 71.1\%$) (low strength of evidence). Additional noninvasive testing was more common following CCTA at the index visit: the larger good-quality trial reported 10.2% vs. 0.9% for SPECT (RD, 9.4; 95% CI, 6.1 to 12.7 per 100 patients) and the smaller fair-quality trial reported 24% vs. 0% for SPECT (RD, 24 per 100 people; $p < 0.001$) (high strength of evidence from 2 trials). Use of additional noninvasive testing through 6 months was similar (1% vs. 3%) (low strength of evidence from 1 trial). Moderate-strength evidence from both trials of ED patients suggested similar referral for revascularization, including PCI and CABG evaluated separately, at the index visit (3.9% vs. 2.1%) and through 6 months (0.5% vs. 0%).

Intermediate to High Pretest Risk of CAD

A total of two comparative studies (in 3 publications) were identified in populations with intermediate to high pretest risk of CAD: PET versus SPECT (1 prospective observational study)[37,38] and CCTA versus SPECT (1 RCT).[39] Table D summarizes the primary findings for this risk category.

The main comparison for which evidence was found is CCTA versus SPECT. One small poor-quality trial of 180 outpatients with intermediate to high risk (65% intermediate and 29% high risk; mean Framingham risk estimate, 18.7) with a mean of 1.8 months followup found no deaths or MIs (insufficient strength of evidence). Strength of evidence was low that cardiac hospitalizations occurred at a similar rate between groups (12% vs. 11%). CCTA was associated with more revascularizations (8% vs. 1%; RD, 6.6%; 95% CI, 0.7% to 12.5%), as well as slightly more ICA referrals (13% vs. 8%; RD, 5; 95% CI, -4 to 14 per 100 people; p not statistically significant) and slightly but not significantly less noninvasive cardiac imaging testing (3% vs. 10%; RD, -7; 95% CI, -14 to 0.4 per 100 people) through the same followup period (low strength of evidence).

High Pretest Risk of CAD

One study in a population with high pretest risk of CAD compared SPECT and exercise ECG.[20] Evidence was based on subgroup analyses and was insufficient for all outcomes. Results are detailed in the full report.

Mixed Population: Pretest Risk Not Reported or Results Not Stratified by Risk

A total of nine comparative studies were identified in populations with mixed pretest risk of CAD or for which risk was not reported. (One administrative database study reported outcomes for 6 different test comparisons.) The study comparisons were CCTA versus usual care (1 RCT),[19] exercise ECG (1 RCT,[40] 1 administrative database[41]), SPECT (1 prospective registry,[42] 1 administrative database[43]), nuclear MPI (1 prospective observational study,[44] 1 administrative database[41]), and stress echocardiography (1 administrative database);[41] SPECT versus exercise ECG (1 RCT,[20] 1 administrative database[41]); and stress echocardiography versus exercise ECG (1 RCT,[45] 1 prospective observational study,[46] 1 administrative database41) and SPECT (1 administrative database).[41] Outcomes with insufficient evidence are not detailed here but are described in the full report. Table E summarizes the primary findings for this risk category.

CCTA Versus Usual Care

In a fair-quality trial of 266 patients presenting to the ED and not stratified by risk (low, 37%; intermediate, 42%; high, 21%), there was low-strength evidence of no difference in 1-month MI (0% vs. 0.8%) or contrast-induced nephropathy (0% in both groups).

SPECT Versus Exercise ECG

In outpatients not stratified by risk (low, 16%; intermediate, 61%; high, 23%), there was low-strength evidence from one fair-quality trial of 457 patients that there was no difference between groups in all-cause mortality (0.8% vs. 0.9%) or MI (0% vs. 0.5%) through a mean of 22 months, while SPECT was associated with fewer revascularizations than exercise ECG (10.8% vs. 17.9%; RD, -7.1; 95% CI, -13.6 to -0.6 per 100 people).

Exercise ECG Versus Nuclear MPI

Low-strength evidence from a large fair-quality administrative database of Medicare outpatients (N = 193,406) suggested that 6-month mortality was

similar between groups (0.78% vs. 1.28%; adjusted odds ratio [OR], 0.93; 95% CI, 0.83 to 1.04). Patients who underwent exercise ECG were less likely to undergo ICA through 6 months than those who were tested with MPI (9.04% vs. 12.13%; adjusted OR, 0.72; 95% CI, 0.70 to 0.75); revascularization, including CABG and PCI evaluated separately, was performed with similar frequency between groups (4.31% vs. 4.59%; adjusted OR, 0.90; 95% CI, 0.85 to 0.94) (low strength of evidence for both).

Stress Echocardiography Versus Nuclear MPI

Low-strength evidence from a large fair-quality administrative database of Medicare outpatients (N = 212,947) suggested that 6-month mortality was similar between groups (0.95% vs. 1.28%; adjusted OR, 1.00; 95% CI, 0.90 to 1.10). Through 6 months, ICA referral was statistically less frequent in the stress echocardiography group (9.50% vs. 12.13%; adjusted OR, 0.78; 95% CI, 0.76 to 0.81), while additional noninvasive testing was slightly more common in this group (5.57% vs. 3.22%; adjusted OR, 1.92; 95% CI, 1.83 to 2.0) (low strength of evidence). There were no apparent clinical differences between groups in referral for revascularization (4.22% vs. 4.59%; adjusted OR, 0.93; 95% CI, 0.88 to 0.98), including CABG and PCI evaluated separately (low strength of evidence).

CCTA Versus Exercise ECG

One fair-quality trial of 500 ED patients not stratified by risk (low, 43%; intermediate, 24%; high, 34%) with 12 months of followup found low-strength evidence of no difference between groups in all-cause mortality (0.4% in both groups) or MI (0.41% vs. 0.82%), while there was moderate-strength evidence that cardiac-related hospitalizations were less common in the CCTA group (0.8% vs. 6.9%; RD, -6.1; 95% CI, -9.5 to -2.7 per 100 people). CCTA was associated with more ICAs (27.2% vs. 20.8%; RD, 6.3; 95% CI, -1.2 to 13.9 per 100 people; p = 0.1011) and more revascularizations (15.2% vs. 7.7%; RD, 7.5; 95% CI, 1.9 to 13.0 per 100 people, including PCI [11.9% vs. 4.9%; RD, 7; 95% CI, 2 to 12 per 100 people]), although CABG was used with similar frequency in both groups (3.3% vs. 2.9%) (low strength of evidence).

CCTA Versus Nuclear MPI

One large fair-quality administrative database study of 141,163 mixed-risk Medicare outpatients provided low-strength evidence that all-cause mortality was similar through 6 months (1.05% vs. 1.28%). CCTA patients were more likely to undergo ICA (22.94% vs. 12.13%; adjusted OR, 2.19; 95% CI, 2.08 to 2.32), additional noninvasive testing (4.98% vs. 3.22%; adjusted OR, 1.52; 95% CI, 1.37 to 1.69), and revascularization (11.41% vs. 4.59%; adjusted OR, 2.76; 95% CI, 2.56 to 2.98), including PCI and CABG evaluated separately, through 6 months (low strength of evidence).

One fair-quality registry study of 1,856 patients provided low-strength evidence that revascularization was more common following CCTA through a median of 1.42 years (% not reported; adjusted OR, 1.62; 95% CI, 1.20 to 2.18); the setting was not reported.

Discussion

Key Findings and Strength of Evidence

Evidence to determine the comparative effectiveness and safety of different noninvasive testing strategies for CAD is limited. While there is a robust body of literature on the diagnostic performance of these tests based on traditional measures of test accuracy (e.g., sensitivity, specificity), only a small number of studies were identified that evaluated the impact of noninvasive testing on clinical outcomes measures in the population of interest for this report. The key findings and strength of evidence for the outcomes identified as being most clinically important are summarized in Tables B–E in the Results section; factors used to determine the overall strength of evidence are summarized in Appendix J of the full report.

A total of 24 comparative studies that evaluated the impact of noninvasive testing on clinical outcomes and/or clinical management outcomes in the population of interest for this report form the basis of this review: 14 RCTs (2 good quality, 9 fair quality, and 3 poor quality)[19-21,25-29,31,32,34,39,40,45] and 10 comparative observational studies (7 fair quality and 3 poor quality).[23,24,30,33,35-38,41-44,46] Common methodological shortcomings in the RCTs include unclear description of randomization sequence and/or test allocation and lack of blinded outcomes assessment. In the observational studies, lack of controlling for confounding and/or blinding of outcomes assessment were common methodological shortcomings. The comparative studies served as the basis of the report and were stratified based on pretest risk, test type (anatomic or functional), and setting. For most outcomes reported in trials, the strength of evidence was rated as low (meaning that our confidence in the estimates of effect is low) based on concerns related to precision and study limitations. However, for some outcomes reported by trials, the strength of evidence

Table B. Summary of findings and strength of evidence: intermediate pretest risk			
Comparison	**Number of Studies (N)**	**Findings***	**Strength of Evidence**
CCTA vs. usual care†	2 RCTs (N = 1,098), 1 observational study (N = 200)	No statistically significant differences between tests were found for all-cause mortality (28–30 days in 2 trials), myocardial infarction (index visit in 1 trial, 28–30 days in 2 trials, and 3 months in 1 observational study), ICA referral (index visit in 1 trial and 28–30 days in 2 trials), any revascularization (index visit in 2 trials), PCI (index visit and 28 days in 1 trial, and 3 months in 1 observational study), CABG (28 days in 1 trial and 3 months in 1 observational study), additional noninvasive testing (index visit and 28 days in 1 RCT), and cardiac hospitalizations (ED index visit in 2 trials, 3 months in 1 observational study).	Low
SPECT vs. exercise ECG	1 RCT (N = 824 women)	SPECT was associated with significantly less additional noninvasive testing (including stress testing with or without imaging) through 24 months (**9.4% vs. 18.6%; RD, -9; 95% CI, -14 to -4 per 100 people**).	Moderate
	2 RCTs (N = 824 women only, N = 280 in intermediate-risk subgroup of a trial of men and women)	No statistically significant differences were found between tests through 24 months for all-cause mortality, revascularization, ICA referral, hospitalization for chest pain (trial of women only), and additional noninvasive testing (trial of the general population). SPECT was associated with a significantly lower referral rate for ICA through 22 months in 1 trial of the general population (**10.6% vs. 43.1%; RD, -32; 95% CI, -43 to -22 per 100 people**).	Low
CCTA vs. functional testing	1 RCT (N = 10,003)	CCTA was associated with a significantly higher referral rate through 3 months for ICA (**12.19% vs. 8.11%; RD, 4.08; 95% CI, 2.90 to 5.26 per 100**), any revascularization (**6.22% vs. 3.16%; RD, 3.07; 95% CI, 2.24 to 3.90 per 100 people**), PCI (**4.8% vs. 2.4%; RD, 2.4; 95% CI, 1.7 to 3.1 per 100 people**), and CABG (**1.44% vs. 0.76%; RD, 0.68; 95% CI, 0.27 to 1.09 per 100 people**).	High
	1 RCT (N = 10,003)	Statistically significant differences between tests were not found for all-cause mortality (12 months and 25 months), nonfatal MI (12 months and 25 months), cardiac hospitalization (25 months), and major procedural complications (stroke, major bleeding). CCTA was associated with a significantly increased risk of hospitalization for unstable angina (1.22% vs. 0.82%; RD, 0.40; 95% CI, 0.01 to 0.80 per 100 people) and minor side effects from testing such as stress-induced symptoms and mild contrast reactions (0.74% vs. 0.42%; RR, 1.77; 95% CI, 1.05 to 3.01), although it is unclear if the differences are clinically meaningful.	Moderate
CCTA vs. SPECT	1 RCT (N = 400)	No statistically significant differences between tests were found for all-cause mortality (median, 24.5 months); cardiac hospitalization (median, 40.4 months); ICA referral, any revascularization, PCI, and additional noninvasive testing, including myocardial perfusion imaging, stress echocardiography, and CCTA (through 12 months); and minor adverse reactions (including headache, nausea, dizziness, or feeling of warmth), rash or pruritus, and post-test renal dysfunction. CCTA was associated with a significantly higher risk of CABG through 12 months (3.5% vs. 0.5%; RD, 3.0; 95% CI, 0.3 to 5.7 per 100 people) and a significantly lower incidence of the composite of periprocedural chest pain, shortness of breath, or palpitations (0.5% vs. 15.9%; RD, -15.4; 95% CI, -20.8 to -10.1 per 100 people).	Low

CABG = coronary artery bypass graft; CCTA = coronary computed tomography angiography; CI = confidence interval; ECG = electrocardiography; ED = emergency department; ICA = invasive coronary angiography; MI = myocardial infarction; PCI = percutaneous coronary intervention; RCT = randomized controlled trial; RD = risk difference; RR = relative risk; SPECT = single-photon emission computed tomography

*Statistically significant differences are indicated in bold font.

†Usual care consisted of a conventional diagnostic strategy using serial ECG and cardiac biomarkers.

Table C. Summary of findings and strength of evidence: low to intermediate pretest risk			
Comparison	**Number of Studies (N)**	**Findings***	**Strength of Evidence**
CCTA vs. usual care†	1 RCT (N = 1,392)	No statistically significant differences between tests were found for cardiac hospitalization after the index visit through 1 month. CCTA was associated with significantly less additional noninvasive testing at the index visit in 1 trial (**13.7% vs. 57.8%; RD, -44.1; 95% CI, -49.2 to -39.1 per 100 people**), as well as through 1 month in 1 trial (**23.1% vs. 66.4%; RD, -43.3; 95% CI, -48.4 to -38.1 per 100 people**). CCTA was also associated with a decreased risk of cardiac hospitalization at the ED index visit in 1 trial (**50% vs. 77%; RD, -26.8; 95% CI, -31.9 to -21.8 per 100 people**).	Moderate
	2 RCTs (N = 1,452), 1 observational study (N = 1,788)	No statistically significant differences between tests were found for all-cause mortality (at 1 month in 1 trial and in the observational study), myocardial infarction (at index visit and 1 month in 1 trial and up to 1 month in the observational study), ICA referral (at index visit in 1 trial and at 1–3 months in 2 trials), any revascularization (at 1–3 months in 2 trials and through 1 month in the observational study), PCI and CABG (both at 3 months in 1 trial), and bradyarrhythmia (1 trial). CCTA was associated with more revascularization procedures at the index visit in 1 trial (**2.5% vs. 0.9%; RD, 1.7; 95% CI, 0.3 to 3.0 per 100 people**), and with less additional stress testing at the index visit in 1 trial (**13.7% vs. 57.8%; RD, -44.1; 95% CI, -49.2 to -39.1 per 100 people**) and through 3 months in the other (**33% vs. 60%; RD, -27; 95% CI, -51 to -2 per 100 people**), as well as through 3 months in 1 observational study (**4% vs. 21%; p <0.001**). ICA referral was less common with CCTA (1% vs. 3%) in the retrospective observational study; although authors reported statistical significance (p <0.001), clinical significance is unclear.	Low
CCTA vs. exercise ECG	1 RCT (N = 562), 1 observational study (N = 498)	No statistically significant differences were found between tests in all-cause mortality (at 30 days in 1 trial and at 12 months in both studies) and myocardial infarction (at index visit and 30 days in 1 trial, and at 12 months in the observational study). CCTA was associated with a significantly higher referral rate for ICA (**9.0% vs. 2.3%; RD, 4.8; 95% CI, 0.8 to 8.9 per 100 people**) and revascularization (**4.3% vs. 1.3%; RD, 3.1; 95% CI, 0.5 to 5.7 per 100 people**) through 12 months in 1 trial.	Low
CCTA vs. SPECT	2 RCTs (N = 952)	CCTA was associated with higher rates of additional noninvasive testing at the index visit: 10.2% vs. 0.9% in the larger trial (**RD, 9.4; 95% CI, 6.1 to 12.7 per 100 patients**) and 24% vs. 0% in the smaller trial (**RD, 24 per 100 people; p <0.001**).	High
	2 RCTs (N = 952)	No statistically significant differences between tests were found for revascularization, PCI, and CABG at the index visit and through 6 months (2 trials), and for cardiac hospitalization through 6 months (1 trial) and 30 months (1 observational study).	Moderate
	2 RCTs (N = 952), 1 observational study (N = 252)	No statistically significant differences between tests were found for all-cause mortality at 6 months (2 trials) and at 30 months (1 observational study), myocardial infarction and ICA referral at index visit and through 6 months (2 trials), and additional testing through 6 months (1 trial).	Low

CABG = coronary artery bypass graft; CCTA = coronary computed tomography angiography; CI = confidence interval; ECG = electrocardiography; ED = emergency department; ICA = invasive coronary angiography; PCI = percutaneous coronary intervention; RCT = randomized controlled trial; RD = risk difference; SPECT = single-photon emission computed tomography

*Statistically significant differences are indicated in bold font.

†Usual care consisted of a conventional diagnostic strategy using serial ECG and cardiac biomarkers.

Table D. Summary of findings and strength of evidence: intermediate to high pretest risk			
Comparison	**Number of Studies (N)**	**Findings***	**Strength of Evidence**
CCTA vs. SPECT	1 RCT (N =180)	No statistically significant differences were found between tests through a mean 1.8 months for ICA referral, additional noninvasive testing, and coronary artery disease–related hospitalization. CCTA was associated with more revascularizations through a mean of 1.8 months (8% vs. 1%; RD, 6.6; 95% CI, 0.7 to 12.5).	Low

CCTA = coronary computed tomography angiography; CI = confidence interval; ICA = invasive coronary angiography; RCT = randomized controlled trial; RD = risk difference; SPECT = single-photon emission computed tomography
*Statistically significant differences are indicated in bold font.

Table E. Summary of findings and strength of evidence: mixed pretest risk			
Comparison	**Number of Studies (N)**	**Findings***	**Strength of Evidence**
CCTA vs. usual care†	1 RCT (N = 266)	No statistically significant differences were found between tests through 30 days for myocardial infarction and contrast-inducted nephropathy.	Low
SPECT vs. exercise ECG	1 RCT (N = 457)	No statistically significant differences were found between tests at 22 months for all-cause mortality and myocardial infarction. SPECT was associated with significantly fewer revascularizations through 22 months (**10.8% vs. 17.9%; RD, -7.1; 95% CI, -13.6 to -0.6 per 100 people**).	Low
Exercise ECG vs. nuclear MPI	1 observational study (N = 193,406 Medicare)	No statistically significant differences between tests were found for all-cause mortality at 6 months. Exercise ECG was associated with significantly fewer referrals for ICA (**9.04% vs. 12.13%; adjusted OR, 0.72; 95% CI, 0.70 to 0.75**), any revascularization (**4.31% vs. 4.59%; adjusted OR, 0.90; 95% CI, 0.85 to 0.94**), and PCI (**2.57% vs. 3.37%; adjusted OR, 0.72; 95% CI, 0.68 to 0.77**), and significantly higher rates of CABG (**1.82% vs. 1.29%; adjusted OR, 1.37; 95% CI, 1.26 to 1.49**) and additional noninvasive testing (**19.34% vs. 3.22%; adjusted OR, 7.46; 95% CI, 7.16 to 7.77**) through 6 months, although it is unclear if the differences for any revascularization, PCI, and CABG are clinically meaningful.	Low

Table E. Summary of findings and strength of evidence: mixed pretest risk (continued)			
Comparison	**Number of Studies (N)**	**Findings***	**Strength of Evidence**
Stress echocardio-graphy vs. nuclear MPI	1 observational study (N = 212,947 Medicare)	No statistically significant differences between tests were found for all-cause mortality at 6 months. Stress echocardiography was associated with significantly fewer referrals for ICA (**9.50% vs. 12.13%; adjusted OR, 0.78; 95% CI, 0.76 to 0.81**), any revascularization (**4.22% vs. 4.59%; adjusted OR, 0.93; 95% CI, 0.88 to 0.98**), and PCI (**2.61% vs. 3.37%; adjusted OR, 0.76; 95% CI, 0.72 to 0.81**), and significantly higher rates of CABG (**1.69% vs. 1.29%; adjusted OR, 1.40; 95% CI, 1.29 to 1.52**) and additional noninvasive testing (**5.57% vs. 3.22%; adjusted OR, 1.92; 95% CI, 1.83 to 2.0**) through 6 months, although it is unclear if the differences for any revascularization, PCI, and CABG are clinically meaningful.	Low
CCTA vs. exercise ECG	1 RCT (N = 500)	CCTA resulted in significantly less additional noninvasive testing (**2.4% vs. 31.3%; RD, -29; 95% CI, -37 to -23 per 100 people**), as well as fewer cardiac rehospitalizations (**0.8% vs. 6.9%; RD, -6.1; 95% CI, -9.5 to -2.7 per 100 people**), through 12 months.	Moderate
	1 RCT (N = 500)	No statistically significant differences between tests were found for all-cause mortality, myocardial infarction, referral for ICA, and CABG through 12 months. CCTA was associated with a significantly increased risk of any revascularization (**15.2% vs. 7.7%; RD, 7.5; 95% CI, 1.9 to 13.0 per 100 people**) and PCI (**11.9% vs. 4.9%; RD, 7; 95% CI, 2 to 12 per 100 people**) through 12 months.	Low
CCTA vs. nuclear MPI	2 observational studies (N = 141,163 Medicare, N = 1,856 general population)	No statistically significant differences between tests were found for all-cause mortality at 6 months in the Medicare population. CCTA was associated with significantly higher referral rates for ICA (**22.94% vs. 12.13%; adjusted OR, 2.19; 95% CI, 2.08 to 2.32**), PCI (**7.85% vs. 3.37%; adjusted OR, 2.49; 95% CI, 2.28 to 2.72**), CABG (**3.71% vs. 1.29%; adjusted OR, 3.00; 95% CI, 2.63 to 3.42**), and additional testing (**4.98% vs. 3.22%; adjusted OR, 1.52; 95% CI, 1.37 to 1.69**) through 6 months in the Medicare population; for any revascularization through 6 months in the Medicare population (**11.41% vs. 4.59%; adjusted OR, 2.76; 95% CI, 2.56 to 2.98**), and for any revascularization through a median 1.42 years in the general population (**% not reported; adjusted OR, 1.62; 95% CI, 1.20 to 2.18**).	Low

CABG = coronary artery bypass graft; CCTA = coronary computed tomography angiography; CI = confidence interval; ECG = electrocardiography; ICA = invasive coronary angiography; MPI = myocardial perfusion imaging; OR = odds ratio; PCI = percutaneous coronary intervention; RCT = randomized controlled trial; RD = risk difference; SPECT = single-photon emission computed tomography

*Statistically significant differences are indicated in bold font.

†Usual care consisted of a conventional diagnostic strategy using serial ECG and cardiac biomarkers.

was found to be moderate or high. For the majority of outcomes reported by comparative observational studies, the strength of evidence was found to be insufficient because of study limitations, although some outcomes were graded as low strength of evidence when the estimates were considered to be at low risk for imprecision and confounding was controlled. Eight RCTs and one observational study were conducted in ED settings or specialized chest pain clinics[40] and compared CCTA with functional testing[31,32,34,40] or usual care.[19,21,23,24,28,29] In these studies, most of the available data were reported for the index ED visit, and with the exception of two trials reporting 12-month followup, the maximum followup in ED studies was 6 months. The remaining 5 trials[20,25,27,39,45] and 13 comparative observational studies were conducted in outpatient, various, or unspecified settings; in general, these studies had longer followup periods, which ranged from a mean of 55 days to 30 months. Pretest risk could not be standardized across studies, and was variably determined and defined across studies. Thus, categories of pretest risk used here are based on how the study authors defined risk.

Clinical Outcomes

There was no clear difference in MI or in all-cause mortality between different testing strategies across settings and pretest risk groups that included patients with intermediate pretest risk, based on low- to moderate-strength evidence from eight trials. The definition of intermediate pretest risk was broad. The frequency of all-cause mortality was low across studies in all settings. In trials enrolling outpatients, the frequency of all-cause mortality ranged from 0 to 1.5 percent for a variety of noninvasive testing strategies, and the frequency in trials in the ED setting past the initial index visit ranged from 0 to 1.08 percent across a variety of noninvasive testing or usual-care strategies, with no statistical difference between any groups. Similarly the frequency of MI was low, ranging from 0 to 0.8 percent (up to a median of 25 months) in outpatient settings and 0 to 3 percent (up to 12 months) in ED settings, with no statistical differences between groups. The strongest evidence came from three trials: one that compared CCTA with functional testing in an outpatient setting[25] and two that compared CCTA with SPECT in an ED setting.[31,32] For the trial of CCTA versus functional testing, which was also the largest trial (N = 10,003), there were no differences in all-cause mortality between groups through 12 months (0.42% vs. 0.64%) or at a median of 25 months followup (1.48% vs. 1.50%) or in nonfatal MI at 12 months (0.36% vs. 0.54%; RD, -0.18; 95% CI, -0.44 to 0.08 per 100 people) or at a median of 25 months followup (0.60% vs. 0.80%; RD, -0.20; 95% CI, -0.53 to 0.13 per 100 people);[25] strength of evidence was moderate for both outcomes. Across the two trials comparing CCTA with SPECT in an ED setting, there was low-strength evidence that there was no difference between tests for mortality or MI; no deaths or MIs were reported through a mean of 6 months past the initial ED visit.[31,32] Across the remaining trials, no difference was found between tests because of lack of precision and study limitations (low strength of evidence). Higher quality observational studies (i.e., those that controlled for confounding) supported these findings. No conclusions can be drawn regarding the impact of testing on clinical outcomes for patients at low risk or high risk (without ECG changes, troponin elevation, or other characteristics of ACS), as only subanalyses of fewer than 100 patients were available.

Several factors may have contributed to finding no statistical differences between tests on clinical outcomes. Given the low incidence of mortality and MI in the studies previously noted, sample sizes in even the largest trials may have been too small to detect differences between tests. The low incidence of mortality and MI suggests that study populations may generally have been at the lower end of the intermediate pretest risk range. Improvements in medical therapy in the past few decades, including use of statins, may contribute to the low incidence of these outcomes. An additional consideration is the possibility that differences between tests in true sensitivity to detect treatable CAD or ability to identify high-risk disease are not large. Small differences in sensitivity may have little impact on the probability of disease when the pretest probability is low. Even if two tests do not have the same sensitivity, the lack of difference in the occurrence of outcome events in most studies between people who were assigned to receive different tests could result from either the lack of efficacy of treatments administered to test-positive people or the lack of difference in the receipt of effective treatments between test-positive and test-negative people. Given that studies do not present data on treatments administered to individual study participants (or how testing directed those decisions), we cannot distinguish between these alternatives. Furthermore, information on post-test risk stratification or treatment based on such stratification was not reported in most studies. Information on clinical decisions and outcomes based on whether tests were positive, negative, or indeterminate was not given in most comparative studies. It is possible that over- or undertreatment may have contributed to similarity in clinical findings. Length of followup may have also impact the findings of no

difference in clinical outcomes. Two larger trials in outpatient settings (SPECT vs. stress ECG[27] and CCTA vs. functional testing[25]) followed patients for 2 or more years. There was insufficient evidence to draw conclusions regarding longer term clinical outcomes from studies in the ED setting because most did not provide data beyond 6 months after the ED visit.

Referral for Invasive Coronary Angiography

There was some variability in conclusions regarding ICA referral following noninvasive testing. In most studies, ICA was more common following CCTA than following various functional tests. The strongest evidence came from one good-quality trial that compared CCTA with functional testing in outpatients; it found that ICA was significantly more common in the CCTA group than the functional testing group by 90 days (12.19% vs. 8.11%; RD, 4.08; 95% CI, 2.90 to 5.26 per 100 people) (high strength of evidence). Interestingly, fewer catheterizations in the CCTA group showed no obstructive CAD (3.4% vs. 4.3%),[25] perhaps because of a lower false-positive rate with CCTA. The strength of the quality of evidence regarding ICA referral was low across the remaining trials. Two fair-quality trials comparing CCTA with exercise ECG suggest that ICA referral is more common following CCTA up to 12 months following an initial ED visit, with RD of 4.8 (95% CI, 0.8 to 8.9 per 100 people) in one trial of patients with low to intermediate risk and RD of 6.3 (95% CI, -1.2 to 13.9 per 100 people) in the trial of mixed-risk patients; statistical significance was not reached and strength of evidence was low because of study limitations and lack of precision.

A large administrative data study in Medicare patients found that ICA was significantly more common following CCTA than following MPI (22.94% vs. 12.13%; adjusted OR, 2.19; 95% CI, 2.08 to 2.32) (low strength of evidence).[41] In contrast, across studies comparing CCTA with usual care, there were no statistical differences between testing strategies in any of the trials regardless of pretest risk or setting. However, in the small high-risk group from one trial, fewer CCTA patients had ICA at the index visit (RD, -18; 95% CI, -37 to 0.8; p = 0.0714) (low strength of evidence). Evidence from observational studies for comparisons of CCTA with other tests was considered insufficient because of study limitations and lack of precision. Regarding comparisons of functional tests, two RCTs[20,27] and one large administrative database study[41] provided low-strength evidence on ICA referral in outpatient settings. One trial comparing SPECT with exercise ECG in intermediate-risk women reported a 6-percent referral for ICA in each test group by 24 months. However, the other trial making this comparison reported a significantly lower frequency of ICA referral by 22 months following SPECT in a subgroup of patients with intermediate pretest risk (RD, -32; 95% CI, -43 to -22 per 100 people), as well as in a subgroup of high-risk patients (RD, -41; 95% CI, -58 to -24 per 100 people)[20] This same trial used Bayesian methods to model post-test risk and reported that 86 percent of those with low pretest risk finished with low post-test risk. Patients in either arm whose tests were normal or indicated low risk did not receive ICA; 3 percent and 38 percent in the intermediate and high post-test risk groups had ICA following SPECT, compared with 13 percent and 85 percent in the intermediate and high post-test risk groups following exercise ECG. This type of modeling is not a standard approach to post-test risk assessment, so the generalizability of these results is not clear. The administrative database study of Medicare patients reported that, compared with nuclear MPI, ICA referral was lower following exercise ECG (OR, 0.72; 95% CI, 0.70 to 0.75) and stress echocardiography (OR, 0.78; 95% CI, 0 .76 to 0.81)[41] (low strength of evidence). Evidence from the remaining observational studies was considered insufficient.

None of the studies provided analysis or explicit information regarding unnecessary treatment or testing.

Revascularization

Findings were inconsistent across diagnostic strategies with regard to revascularization referral. There was high-strength evidence from one large trial that any revascularization within 90 days was more common following CCTA compared with functional testing (RD, 3.07; 95% CI, 2.24 to 3.90 per 100 patients); the same was true for PCI specifically (RD, 2.4; 95% CI, 1.7 to 3.1 per 100 patients)[25] (high strength of evidence). Revascularization was also more common 6 to 12 months following CCTA compared with exercise ECG across two studies (1 RCT, 1 observational study)[40,41] of mixed-risk ED patients (low strength of evidence), as well as across two observational studies comparing CCTA with nuclear MPI[41,44] in outpatient settings up to 1.4 years (low strength of evidence). In contrast, the frequency of revascularization was similar for CCTA and SPECT (pooled RD, 2 per 100 patients; 95% CI, 0 to 4 per 100 patients) at the index ED visit and at 6 months (pooled RD, 0; 95% CI, 0 to 1 per 100 patients) across two trials (moderate strength of evidence).[31,32] PCI and CABG frequencies in these trials were also similar between tests; strength of evidence was moderate. Further, there was low-strength evidence of no statistical differences

in revascularization frequency between CCTA and usual care at the index visit or at 1 to 3 months followup based on data from four trials.[19,21,28,29] Evidence comparing functional tests was inconsistent, with one small trial reporting fewer revascularizations following SPECT than exercise ECG (RD, -7.1; 95% CI, -13.6 to -0.6 per 100 people)[20] (low strength of evidence) and one large Medicare administrative database study reporting a similar frequency of revascularization, including PCI and CABG, for exercise ECG (4.31% vs. 4.59%) and stress echocardiography (4.22% vs. 4.59) as for nuclear MPI (low strength of evidence). For the latter study, although the differences between groups were statistically significant for both comparators, they may not be clinically significant. Studies did not describe post-test reclassification of risk or decisionmaking for treatment.

Additional Noninvasive Testing

Additional noninvasive testing, which impacts the cost and efficiency of care, was common in most studies. In the ED setting, there was high-strength evidence from two trials of patients with low to intermediate risk that additional noninvasive testing was significantly more common following CCTA than SPECT at the index visit (RD for largest trial, 9.4; 95% CI, 6.1 to 12.7 per 100 patients).[31,32] In the same setting, there was moderately strong evidence that CCTA was associated with less frequent noninvasive testing compared with usual care at the index visit in one trial[28] and compared with exercise ECG through 12 months past the index ED visit[40] in another trial. In intermediate-risk patients, the frequency of additional testing following CCTA was similar to the frequency following usual care up to 1 month past the ED visit in one trial (low strength of evidence), possibly because many in the usual-care group also received noninvasive imaging.[21] In outpatient settings, the strength of evidence was moderate that SPECT was associated with significantly less additional noninvasive testing compared with exercise ECG through 22 months, based on one large trial of intermediate-risk women (RD, -9; 95% CI, -14 to -4 per 100 people),[27] as well as a from a subgroup of intermediate-risk patients in another trial (RD, -38; 95% CI, -48 to -29 per 100 people).[20] These results likely indicate greater clinician confidence when stress testing is paired with imaging, based on general understanding from accuracy studies that positive and negative predictive values are better for SPECT than for stress testing. In the Medicare administrative database study, both CCTA and stress echocardiography were associated with a significantly higher frequency of additional noninvasive testing compared with nuclear MPI (OR, 1.52; 95% CI, 1.37 to 1.69 and OR, 1.92; 95% CI, 1.83 to 2.0, respectively), but strength of evidence is low. Studies generally did not describe post-test reclassification of risk or decisionmaking for related further testing.

Hospitalization

Cardiovascular-related hospitalizations varied somewhat among pretest risk groups across studies. There was moderate-strength evidence from one large trial of ED patients with low to intermediate risk that the CCTA group was significantly less likely than the usual-care group to be hospitalized or admitted for observation at the index visit (RD, -26.8; 95% CI, -31.9 to -21.8 per 100 people), but that after this visit through 1 month, there was no difference (3% for CCTA vs. 2% for usual care).[28] Low-strength evidence from a large trial of intermediate-risk ED patients suggested that there were fewer hospitalizations following CCTA compared with usual care at the index visit (RD, -33; 95% CI, -39 to -28 per 100 patients).[21] These data imply clinician confidence in the negative predictive value of the anatomic test, yet there is a predisposition of patients to return with unexplained symptoms that can be from a variety of other causes of chest pain, including vasospasm and microvascular dysfunction. In contrast, no statistical differences between CCTA and usual care were identified for ACS hospitalization at the index visit based on subgroups of low- or high-risk patients in one trial,[19] but strength of evidence was low. There was moderate-strength evidence that there was no difference in cardiovascular hospitalizations between CCTA and functional testing groups in low- to intermediate-risk ED patients within 6 months (0% in both groups) based on one trial[31] and through 30 months based on one observational study[33] that compared CCTA with SPECT. In another trial of mixed pretest risk patients presenting to specialized chest pain clinics,[40] moderate-strength evidence suggested that hospitalization for cardiac causes occurred less frequently in the CCTA group compared with the exercise ECG group (RD, -6.1; 95% CI, -9.5 to -2.7 per 100 people) through 12 months. Two trials conducted in outpatient settings reported no differences in cardiac-related hospitalizations between groups. The strongest evidence came from the large trial comparing CCTA with functional testing, which reported no differences at a median of 25 months (RD, -0.30; 95% CI, -0.10 to 0.71 per 100 people)[25] (moderate strength of evidence). The trial of SPECT versus exercise ECG in women also found no difference between groups (low strength of evidence).[27]

Special Populations

With regard to evaluation of special populations, one good-quality trial comparing CCTA with functional testing reported that none of the prespecified subgroups modified the primary composite outcome (all-cause death; nonfatal MI; hospitalization for unstable angina; or a major procedural complication, such as stroke, major bleeding, anaphylaxis, or renal failure requiring dialysis). Results across subgroups were consistent with those for the entire study population. Subgroups examined included age sex, race, pretest risk assessment, CAD equivalence, and pretest probability of CAD.[25] None of the other studies identified evaluated differential effectiveness or safety for the primary outcomes. As noted earlier, one fair-quality trial of exercise SPECT compared with exercise ECG in women found no differences between tests for mortality, ICA referral, revascularization, or hospitalization, but that trial reported a significantly lower use of additional noninvasive testing following SPECT.[27] The strength of evidence was moderate for additional testing and low for other outcomes. An additional small poor-quality RCT in women compared stress echocardiography with exercise ECG; this trial reported similar frequency of a composite outcome that included cardiac death, MI, unstable angina, or coronary angiography demonstrating 50-percent or more luminal narrowing (7.7% vs. 7.4%).[45] However, the strength of evidence was insufficient because of high risk of bias, lack of precision, and unknown consistency. Also as noted earlier, a large fair-quality administrative data study in the Medicare population was identified.[41] Consistent with findings in other studies, this study found no differences in adjusted effect estimates for all-cause mortality for the comparisons of nuclear MPI with stress echocardiography, exercise ECG, or CCTA. CCTA was significantly associated with increased referral for ICA and revascularization (particularly PCI) and use of additional noninvasive testing compared with nuclear MPI (strength of evidence was low for these outcomes and comparisons).

Harms and Consequences of Testing

Harms of testing were rarely reported and details on comparisons of harms for tests were sparse, with many studies stating only that no harms were observed and not providing further detail; 16 of the 27 comparative studies made no mention of evaluation of harms. There were no compelling safety outcomes data that can be used to recommend one approach versus another (low or insufficient strength of evidence). No differences in major procedural complications were identified in the trial comparing CCTA with functional imaging, although mild contrast reactions were significantly more common in the CCTA group than in the functional testing group (moderate strength of evidence).[25] No differences were reported between CCTA and usual care in bradyarrhythmia in one trial[28] or periprocedural complications in another[21] (low strength of evidence for both). A third trial reported that there was no clinical or laboratory evidence of contrast-induced nephropathy in either the CCTA or the usual-care group.[19] One observational study reported incidental findings requiring further investigation in 7.1 percent of those receiving CCTA (insufficient evidence).[33] Evidence from observational studies regarding test-related harms and impact of incidental findings following CCTA was insufficient to draw conclusions.

An important patient safety concern related to noninvasive testing is exposure to low to moderate levels of ionizing radiation, which add to cumulative lifetime radiation exposure. To the extent that noninvasive tests for CAD reduce the need for conventional angiography, cumulative exposure might be reduced. To the extent that they result in the need for additional testing, it may be increased. The true attributable risk from radiation-based diagnostic tests cannot be determined. Some experts consider the potential for harm from radiation exposure (based on either deterministic or stochastic modeling) to be clinically significant, particularly since patients may be likely to have additional tests using radiation over many years. Estimates of radiation exposure from included studies are provided in Appendix G of the full report (Table G4); the Introduction section of the full report provides contextual information on radiation exposure ranges for testing. Radiation exposure from included studies for initial t esting strategies ranged from 3.8 to 17 mSv for CCTA and 10.5 to 38 for SPECT. One study reported a mean of 4.0 mSv for PET,[38] and another study[21] reported a mean of 4.7 mSv for usual care. Consideration of cumulative radiation exposure related to downstream testing and intervention is important when discussing with patients the benefits and consequences of the different noninvasive tests and their contribution to lifetime radiation exposure. Higher mean cumulative radiation accounted for by additional testing was seen in single trials following CCTA compared with usual care (14.3 ± 10.9 vs. 5.3 ± 9.6 mSv)[21] and functional testing (12.0 ± 8.5 vs. 10.1 ± 9.0 mSv).[25] One study reported higher cumulative exposure following CCTA than following SPECT in patients referred for ICA (median, 15.2 mSv; interquartile range, 12.7 to 17.1 vs. median, 10.8 mSv; interquartile range, 10.2 to 11.7).[42] In contrast, another trial reported lower cumulative exposure for additional testing following CCTA versus SPECT (median, 7.3 mSv; interquartile range, 5.1 to 13.7 vs. median, 13.3 mSv; interquartile range, 13.1 to 38.0).[39] One

observational study of CCTA and exercise ECG reported greater cumulative radiation exposure as a result of index plus downstream testing for CCTA in patients whose tests were negative, positive, or inconclusive. However, among those who tested positive and had revascularization, mean cumulative exposure was slightly higher in the ECG group (28 vs. 32 mSv).[35] Consideration of patient preferences with regard to the impact of radiation exposure should be part of shared decisionmaking around noninvasive testing.

Findings in Relationship to What Is Already Known

Few prior reviews have evaluated the impact of noninvasive testing on clinical and management outcomes. Systematic reviews and studies on noninvasive testing for CAD identified from our search focused on traditional measures of test performance (e.g., sensitivity, specificity) compared with ICA and generally did not directly compare the effectiveness and safety of different modalities with regard to impact on clinical outcomes specifically in the population of interest in this report. Consistent with this review, prior systematic reviews[47,48] have reported few or no comparative studies evaluating the impact of noninvasive tests on clinical outcomes, decisionmaking, or use of additional testing, and they note that harms are rarely reported. Relevant studies from these reports were included in this systematic review. The recent AHRQ report on noninvasive testing for CAD in women reported that there was insufficient evidence from three studies that treatment decisionmaking and clinical outcomes were impacted by noninvasive testing;[49] consistent with our report, there were no differences in clinical events or hospitalization in studies comparing noninvasive tests. The authors also concluded that studies were underpowered to detect clinical outcomes.

Applicability

A number of factors that impact the applicability of this report's findings are discussed in this section.

Patients

Eight of the 13 trials identified were in patients presenting to the ED with CAD symptoms; however, the largest trial was in an outpatient setting. Patients presenting to the ED represent a broad spectrum of pretest risk probabilities, including those at low or intermediate risk as well as those at high risk for CAD. The severity, newness, and duration of symptoms may differ from those seen in outpatient settings, where patients generally present with more mild to moderate symptoms. Definitions of pretest risk varied across included studies, and some did not report or stratify by pretest risk, making it difficult to fully evaluate results based on pretest risk across settings. It is likely that the patients enrolled in the included studies are representative of those in the broad range of clinical practice regardless of setting.

Interventions and Comparators

The evidence may be skewed toward newer testing modalities, and studies of established tests may not reflect current technology and diagnostic performance. CCTA was the noninvasive test most often assessed, accounting for 48 percent of included studies. The high proportion of studies dealing with CCTA may be because it is a newer modality and thus is compared with established tests, such as stress echocardiography and MPI. Few studies comparing different types of functional testing, particularly established functional tests, such as stress echocardiography, exercise ECG, and nuclear stress testing, were identified. A recent systematic review suggests that over the past 2 decades, there has been a substantial decline in investigations related to echocardiography and nuclear cardiology, compared with a marked increase in cardiac CT imaging studies.[50] Input from clinical team members and the Technical Expert Panel suggests that there is substantial variation in clinical practice with regard to which test may be ordered as an initial test based on patient presentation, testing availability, and clinical perspective. The applicability of this report may be impacted by lack of clarity on the extent to which CCTA may or may not be the initial noninvasive test for firstline evaluation of symptomatic patients without known CAD after a resting ECG. None of the included studies included a "no testing" arm. To the extent that clinical decisionmaking is based on clinical evaluation and judgment without testing, findings in this report may be less applicable to settings where testing is not routinely done.

Outcomes

Findings related to rare outcomes of death, MI, or hospitalization may not be fully applicable to broader clinical populations, in part because of small study sizes and inability to fully characterize such outcomes, particularly over the longer term. Moreover, the impact of a negative test or the treatment downstream from a positive test may extend beyond traditional major adverse coronary events to quality of life, reduction in symptoms, and level of activity. These outcomes were not examined in the majority of included studies. The majority of trials reported outcome at the time of an index ED visit, and the

clinical management objectives are somewhat different in an ED setting than in an outpatient setting.

Settings

Most RCTs were conducted in the ED, where test data help determine immediate disposition for discharge or the need for additional evaluation and/or hospitalization. The initial goal is to make a diagnosis for the cause of chest pain in order to inform appropriate treatment and next steps at the index visit. Thus, MI reported at the index visit may reflect a test's ability to make the diagnosis for immediate decisionmaking but not the test's ability to impact future clinical outcomes. Testing is able to affect events only after the index visit, and long-term followup from ED studies was limited. Thus the applicability of findings from ED studies to general outpatient settings over the long term is likely limited. Six RCTs evaluating CCTA were multicenter studies; five were in single-center sites. It is possible that results from single-center trials may be different and less generalizable than results from multicenter trials. Assessing discernible patterns between the multicenter and single-center site studies in this report is a challenge given the heterogeneity across studies with regard to pretest risk and how comparators such as usual care are defined.

Implications for Clinical and Policy Decisionmaking

The 2012 ACCF/AHA Guideline states that diagnostic testing is most valuable when the pretest probability of ischemic heart diseases is intermediate (10%–90%) and provides a range of options for which test may be used in a given scenario.[9] However, the effectiveness of different modalities with regard to impact on clinical outcomes is not compared. Currently, a variety of tests as the initial (and additional) diagnostic tests for patients at intermediate pretest risk of CAD are employed, and there is uncertainty regarding which tests, if any, may be most suitable and beneficial in patients who present with symptoms suggestive of CAD but have no prior history of it. Although several ACCF/AHA Appropriate Use Criteria are available, including the 2013 multimodality imaging Appropriate Use Criteria,[51] they do not explicitly compare multiple noninvasive testing modalities, nor do they make specific recommendations for the timing and sequencing of tests or for repeat testing based on pretest risk group.

Low- to moderate-strength evidence from nine trials suggested that there is no clear difference in MI or in all-cause mortality between different testing strategies across settings and pretest risk groupings that included those at intermediate risk. Possible contributors to this finding, including lack of power to detect a difference, were previously described. Information from two studies that provided data on groups with low and high pretest risk (without ACS) do not provide insight into the best testing strategies in those groups; the strength of evidence was insufficient for the few outcomes reported and no conclusions can be drawn. Across studies that enrolled intermediate-risk groups, no clear benefits of one testing strategy versus another were seen, and no clear picture of harms for various tests was available from included studies. One apparent trend uncovered by the review is that tests that evaluate coronary anatomy, such as CT, result in a greater likelihood of referral for ICA and subsequent intervention than functional tests do; however, the strength of evidence varied from high to low depending on the comparator, and the impact on clinical outcomes is not known, as most studies did not present data on treatments administered to individual study participants. Thus, it is not clear if the increased referrals were helpful or not with regard to influencing clinical outcomes. In addition, potential harm from use of invasive treatments (which carry specific risks) if clinical benefit is not clear was not described. Only two studies provided limited information on the overall impact of testing and resulting treatment strategies on patient symptoms and quality of life. No studies that compared testing with an arm that received no testing were identified, so the impact of any of the noninvasive testing pathways on clinical evaluation is not known.

As defined in the ACCF/AHA Guideline, the intermediate pretest group is broad and heterogeneous (10%–90%), and in the absence of information on post-test risk, the value of the various tests for influencing important management decisions at each end of the spectrum is not clear. The ACCF/AHA Guideline and various Appropriate Use Criteria[52-55] provide general recommendations for testing and treatment.

In general, next steps following a positive result from an initial noninvasive test are in part based on the post-test annual predicted rate of cardiac mortality as described in the 2012 ACCF/AHA Guideline: low risk (<1% per year), intermediate risk (1%–3% per year), or high risk for cardiac mortality (>3% per year).[9] Clinical presentation and test results are both considered in this determination. In general, for people who would be categorized as being at low risk (negative test result) or intermediate risk and who do not exhibit characteristics of ACS, medical management may be appropriate. In most instances, patients in these categories can be managed without

invasive assessment. In patients who are considered to be at high risk based on noninvasive testing and presentation, ICA for further risk stratification and assessment of appropriateness for revascularization may be the next logical steps. In general, indications for revascularization are based on the clinical presentation (ACS or stable angina); the severity of the angina (based on Canadian Cardiovascular Society Classification); the extent of ischemia on noninvasive testing; and the presence or absence of other prognostic factors, including congestive heart failure, depressed left ventricular function, and diabetes; the extent of medical therapy; and the extent of anatomic disease.[56,57]

Thus, post-test disease probability is an important factor in determining next steps for testing and treatment. From the included studies, however, it is not clear how post-test risk was assessed, which clinical pathways were followed after the initial test, which test(s) may lead to the most appropriate treatment given the post-test risk, or whether the treatments impacted outcomes. While the ACCF/AHA Guideline and various Appropriate Use Criteria provide a range of options for which test may be used in a given scenario and which treatment initiated, the effectiveness of different testing modalities leading to appropriate treatment are not compared with regard to impact on clinical outcomes.

In the absence of high-strength evidence regarding testing options, including the possibility of not testing, decisions must necessarily be made on the basis of other factors related to the initial test and potential followup. The ability of a test to accurately diagnose treatable CAD is important; so too are the costs and consequences beyond the initial test, such as followup of false-negative results (e.g., tests with high false-positive rates in a population with low pretest risk), and the costs and consequences of missing significant disease (e.g., dismissal from the ED of patients with CAD needing treatment). The costs and consequences depend to some extent on the role a test plays in the diagnostic workup pathway, as well as the availability and convenience of a test. Patient pretest probability of disease and consideration of the likelihood ratios with regard to goals of ruling in or ruling out CAD should be a part of the decisionmaking process. Consequences of testing that need to be considered include those related to patient anxiety and patient quality of life and those related to radiation exposure of the index test, as well as potential downstream exposure from additional testing resulting from the initial test and future testing and/or treatment. Consideration of patients' preferences based on their understanding the range of consequences of initial and downstream testing is an important part of shared decisionmaking for initiating noninvasive testing.

Limitations of the Systematic Review Process

This review has some potential limitations. Stratifying by pretest risk, which was in keeping with the intent of the Key Questions, may have resulted in fewer studies to pool and left single studies for most comparisons. This, combined with substantial heterogeneity in how pretest risk was defined, the timeframes over which outcomes were evaluated, and clinical heterogeneity between the tests evaluated, resulted in too few studies for head-to-head meta-analysis for most outcomes, and network meta-analysis was not feasible.

Variable reporting on patient symptoms and characteristics related to CAD risk precluded application of a standardized method for calculating or assigning pretest risk across studies. In light of this, test comparisons were evaluated according to pretest risk as specified by authors to discern patterns within and across pretest risk levels and settings, and qualitatively synthesize outcomes when pooling was not possible. This approach resulted in limited ability to truly examine the evidence by pretest risk.

Inclusion was restricted to studies published in English; however, this is not likely to have impacted the evidence base, as few potential non–English-language studies were seen in the searches. Given the paucity of RCTs, comparative observational studies were included. Despite a focus on outcomes in studies that controlled for confounding, there is a possibility that residual confounding influenced reported results, lowering confidence in effect estimates. The comparative studies included may not adequately capture harms safety issues in the population of interest. The focused criteria on inclusion of studies comparing an established firstline test (beyond a resting ECG) narrowed the review scope substantially, but this focus was intended to provide a clearer approach to addressing the areas of uncertainty. It is possible that older historical studies outside of our population of interest could provide more detailed information about the safety of various tests, particularly more established tests.

There were too few studies of any given comparison to meaningfully evaluate reporting and publication bias. Where available, protocols of trials were reviewed to consider the extent to which outcomes were reported selectively, and information from Scientific Information Packets requested from stakeholders was evaluated; while overt publication bias was not detected, there is always the possibility it may be present. This review provides a snapshot of currently available evidence on the questions

posed. Included studies may not reflect technological advances that have been made in the various testing modalities.

Limitations of the Evidence Base

Important limitations of the evidence base include the paucity of studies that compared the impact of different noninvasive tests on hard clinical outcomes, such as mortality and myocardial infarction; few RCTs were available, in particular for comparisons of established functional tests in the population of interest. No trials that included a no-testing arm were identified. Methods for assessing pretest risk, defining cardiovascular outcomes, and defining usual care were poorly reported and not standardized. The variable methods for determination and classification of pretest risk across studies and inability to implement a standardized method for assessing pretest risk across studies precluded detailed evaluation of testing strategies by pretest risk level to determine the comparative values of tests for a given pretest risk. The intermediate risk range is broad (10%–90%). Studies did not provide information on the impact of test results on post-test risk stratification or clinical decisionmaking for treatment or further testing, precluding evaluation of the impact of testing in this group. Some studies reported composite cardiovascular outcomes, which can be misleading, depending on the effects on the individual components.[58] Studies did not evaluate aspects of unnecessary testing. Reporting of harms was suboptimal; 16 of the 27 comparative studies made no mention of evaluation of harms and another 3 merely stated that there were no adverse events. With the exception of one study, authors reported few details about harms. As mentioned previously, study sample sizes and short-term followup may preclude evaluation of rare events. Studies did not describe the impact of testing on treatment choices. Few studies on PET, CACS, and established tests such as stress echocardiography were identified.

Research Gaps and Recommendations

The gaps in the available evidence are many. Two primary issues relate to the need to improve reporting and standardization of pretest CAD risk and to enhance the evidence linking testing strategies and clinical pathways with clinical outcomes. Use of standardized risk models that refine and narrow the currently broad "intermediate–risk" group is needed. For example, because of health care trends to streamline and reduce the cost of care, newer risk models such as the Duke Clinical Score have narrowed the intermediate range and tend to reclassify many of those classified as "intermediate risk" in the Diamond and Forrester model to "low risk."[59] Documentation of post-test risk stratification and its impact on clinical management (treatment and referral for additional testing) is needed to determine optimal testing strategies and roles of tests in different pretest risk groups. This may facilitate comparison of tests to effectively parse out patients at the highest risk end and those at the lower risk end, as well as evaluation of the impact of management decisions in these groups, as they likely will differ. Documentation of management of those who test positive compared with those who test negative and followup of these groups for sufficient time to evaluate clinical outcomes are needed. Prospective cohort studies that address selection bias and confounding by indication have the potential to enhance the evidence base and may be more feasible than RCTs for some settings. Studies comparing testing versus clinical evaluation without testing would provide valuable information for assessing the need for testing, possible overuse of testing, and the impact of testing in general. Comparative studies (RCTs, pragmatic trials, or prospective cohorts) of functional tests that reflect technological advances as applied to symptomatic patients without known CAD would update the evidence base. Meta-analysis of patient-level data from existing trials may allow for more specific stratification by pretest probability or specific risk factors. Important insights into the overall impact of testing on long-term outcomes could come from studies that (1) document how test results specifically influence decisionmaking regarding further testing and treatment strategies, and (2) follow patients to evaluate the impact of the testing pathway. Future research also needs to incorporate evaluation of patient-centered outcomes, such as quality of life, symptom status, and the impact of testing.

Primary gaps and considerations for future research are summarized in Table F.

Conclusion

A review of current studies found no clear differences between testing strategies across settings with regard to clinical or management outcomes that would allow recommendation of one strategy over another for any given pretest risk group that included patients with intermediate pretest risk. No conclusions regarding low-risk patients or those without ACS at high risk are possible. Limited evidence from RCTs found no clear differences between CCTA versus other strategies in clinical outcomes across risk groups, although anatomic testing may result in a higher frequency of referral for ICA and revascularization. The frequency of all-cause

Table F. Overview of research gaps and recommendations		
Research Components	**Evidence Gap**	**Future Research Recommendations**
Study design methods and reporting	Gaps include lack of a standardized approach to determining and reporting pretest risk across studies; variable definitions of pretest risk, which precluded effective stratification by pretest risk; the large range of pretest likelihoods for "intermediate" risk patients (10%–90%), which precluded detailed evaluation of the impact of testing for patients at the lowest and highest ends of the range.	A standardized approach for determination of pretest risk that can be applied across study designs is needed. Future research should use risk models that further refine the range of pretest probability for those at intermediate risk (e.g., the Duke Clinical Score) to delineate the impact of testing on clinical decisionmaking at the lower and higher ends of the range. Tools that refine the range may also be clinically useful.
	Studies describing outcomes at the index ED visit do not allow conclusions regarding the impact of testing on clinical outcomes over the longer term.	Longer followup (>12 months) and documentation of the impact of testing on treatment decisions and hard clinical outcomes are needed. RCTs, pragmatic trials, or prospective cohort studies that address selection bias and confounding by indication could be employed.
	None of the included studies evaluated issues of unnecessary testing or treatment in patients without known CAD.	As a first step, a priori definitions for necessary vs. unnecessary testing or treatment are needed, and they should be evidence based. Given the variability of clinical practice and medicolegal concerns, this may be challenging. Evaluation of Appropriate Use Criteria and examination of evidence on the clinical outcomes based on application of such criteria may help further define necessary vs. unnecessary.
Patient populations	There is a paucity of studies on patients with low or very low pretest probability of CAD, and the value of testing is not clear for this population.	Studies (RCTs, pragmatic trials, or methodologically rigorous comparative cohort studies) that compare a testing strategy (and related clinical management) with a strategy of no testing (and related clinical management) are needed. Sufficient sample size may be a challenge, given the low prevalence of CAD that is likely in this group.
	Few active trials listed in ClinicalTrials.gov pertain to symptomatic patients without known CAD, yet this group of patients commonly presents for evaluation and testing, particularly in outpatient settings. (See Appendix K in the full report.)	Future studies focused on those without known/prior CAD history or studies that analyze outcomes for this group of patients separately from those with known CAD are needed.
	There is a paucity of high-quality studies comparing various testing strategies in outpatient clinic populations.	Studies of patients who typically present in outpatient settings are needed. Greater integration of cardiologists into hospital settings may facilitate the conduct of studies of outpatients and enhance opportunities for followup of patients initially presenting to the ED.
	Studies do not generally report the extent to which clinical decisionmaking and clinical outcomes may be modified by patient characteristics, sociodemographic factors (e.g., age, sex, race, ethnicity, education, socioeconomic status), or provider characteristics.	RCTs or pragmatic trials with sufficient sample size to compare differential effectiveness and safety of testing strategies based on prespecified analyses are needed.

Table F. Overview of research gaps and recommendations (continued)		
Research Components	**Evidence Gap**	**Future Research Recommendations**
Interventions and comparators	There is a lack of studies comparing outcomes following testing and resulting treatment strategies vs. a strategy of clinical evaluation without testing and resultant treatment strategies.	Studies (RCTs, pragmatic trials, or methodologically rigorous comparative cohort studies) that compare a testing strategy (and related clinical management) with a strategy of no testing (and related clinical management) are needed.
	Older studies of established tests (particularly functional tests) may not be as applicable in light of advances in technology. There was a paucity of studies comparing functional tests with each other.	Studies (RCTs, pragmatic trials, or methodologically rigorous comparative cohort studies) that compare functional tests using more state-of-the art technology and methods with each other and with anatomic tests are needed. New studies should focus on the impact each test makes on clinical decisionmaking and hard clinical outcomes.
Outcome measures	Studies comparing of the impact of noninvasive testing on hard clinical outcomes in those without known CAD are few compared with studies of test accuracy.	Additional sufficiently powered studies examining the impact of testing on hard clinical outcomes (death, MI) at longer term followup (>12 months) are needed.
	There is limited high-quality comparative evidence linking established tests with clinical decisionmaking and subsequent outcomes in the population of interest by pretest risk, particularly in nonemergent settings and over the longer term. Further, there is limited evidence on the impact of tests on post-test risk stratification and the best testing strategy(ies) for post-test risk stratification to identify patients who may be at highest risk and may benefit most from various treatment strategies. It is not clear whether the individuals who would most benefit from given treatment strategies were referred to those strategies and whether the strategies were effective.	Studies that document and compare tests with regard to their impact on prespecified clinical decisionmaking components (e.g., referral for additional testing, initiation or change in medication), particularly in outpatient settings, are needed. Such documentation should also include post-test risk stratification and factors that influenced its determination, what decisions were made based on the test results (positive, negative, or inconclusive results), and impact on hard clinical outcomes (death, MI) over time.
	There is limited evidence on the impact of testing strategies (including consequences of downstream testing and treatment) on patient-related outcomes, such as quality of life and symptom status.	Future studies should incorporate standardized validated measures for patient-reported outcomes and document the impact of testing, including downstream testing, on patient psychological status (particularly with false-positive results), health status, and resource use.
	Adverse events and consequences of testing are poorly reported.	Future study protocols should delineate, a priori, possible adverse events and consequences (including those related to psychological aspects of testing, radiation exposure, resource use) and report their occurrence per the protocol.
Analysis	The lack of a standardized approach to determining and reporting pretest risk across studies and variable definitions of pretest risk used in included studies precluded the ability to effectively stratify by pretest risk or pool data.	Individual patient data meta-analysis of RCTs may provide opportunities to use a standardized approach for pretest risk stratification and may facilitate evaluation of modification by patient characteristics and other factors.
	A number of studies did not provide details for pretest risk or report results stratified by pretest risk.	Studies should stratify by pretest risk of CAD using a standard method and report outcomes based on pretest risk strata.

CAD = coronary artery disease; ED = emergency department; MI = myocardial infarction; RCT = randomized controlled trial

mortality and MI was low across studies in all settings. The absence of information on post-test risk stratification and subsequent decisionmaking precluded evaluation of the impact of testing on patient management or outcomes of management. Testing strategies vary in radiation exposure; there is inadequate comparative evidence to make judgments regarding exposure for the initial test or downstream testing. Assessment of harms was limited. Future research using more refined evidence-based definitions of pretest risk, coupled with information on post-test risk stratification, its impact on clinical management (treatment and referral for additional testing), and longer term followup to assess clinical outcomes, is needed to determine optimal testing strategies and roles of tests in different pretest risk groups.

References

1. Zhu Y, Pandya BJ, Choi HK. Prevalence of gout and hyperuricemia in the US general population: the National Health and Nutrition Examination Survey 2007-2008. Arthritis Rheum. 2011 Oct;63(10):3136-41. PMID: 21800283.

1. 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. PMID: 23245604.

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. PMID: 24352519.

3. Ruigomez A, Rodriguez LA, Wallander MA, et al. Chest pain in general practice: incidence, comorbidity and mortality. Fam Pract. 2006;23(2):167-74. PMID: 16461444.

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. PMID: 21262990.

5. Detry JM. The pathophysiology of myocardial ischaemia. Eur Heart J. 1996 Dec;17 Suppl G:48-52. PMID: 8960456.

6. Schoen FJ, Mitchell RN. Chapter 13, The heart. In: Kumar V, Abbas AK, Aster JC, eds. Robbins & Cotran Pathophysiologic Basis of Disease. 9th ed. Philadelphia: W. B. Saunders Company; 2015. p.524-78.

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. Wolk MJ, Bailey SR, Doherty JU, et al. ACCF/AHA/ASE/ASNC/HFSA/HRS/SCAI/SCCT/SCMR/STS 2013 multimodality appropriate use criteria for the detection and risk assessment of stable ischemic heart disease: a report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Heart Association, American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Failure Society of America, Heart Rhythm Society, Society for Cardiovascular Angiography and Interventions, Society of Cardiovascular Computed Tomography, Society for Cardiovascular Magnetic Resonance, and Society of Thoracic Surgeons. J Am Coll Cardiol. 2014 Feb;63(4):380-406. PMID: 24355759.

9. 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. PMID: 23182125.

10. Cheng VY, Berman DS, Rozanski A, et al. Performance of the traditional age, sex, and angina typicality-based approach for estimating pretest probability of angiographically significant coronary artery disease in patients undergoing coronary computed tomographic angiography: results from the multinational coronary CT angiography evaluation for clinical outcomes: an international multicenter registry (CONFIRM). Circulation. 2011 Nov;124(22):2423-32, 1-8. PMID: 22025600.

11. 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.

12. Higgins JPT, Green S, eds. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. [Updated March 2011.] The Cochrane Collaboration; 2011. www.cochrane-handbook.org.

13. Owens DK, Lohr KN, Atkins D, et al. AHRQ series paper 5: grading the strength of a body of evidence when comparing medical interventions--Agency for Healthcare Research and Quality and the Effective Health-Care Program. J Clin Epidemiol. 2010 May;63(5):513-23. PMID: 19595577.

14. Lu G, Ades AE. Combination of direct and indirect evidence in mixed treatment comparisons. Stat Med. 2004 Oct 30;23(20):3105-24. PMID: 15449338.

15. Hardy RJ, Thompson SG. A likelihood approach to meta-analysis with random effects. Stat Med. 1996 Mar 30;15(6):619-29. PMID: 8731004.

16. Higgins JP, Thompson SG, Deeks JJ, et al. Measuring inconsistency in meta-analyses. BMJ. 2003 Sep 6;327(7414):557-60. PMID: 12958120.

17. 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. Accessed May 14, 2015.

18. Atkins D, Chang SM, Gartlehner G, et al. Assessing applicability when comparing medical interventions: AHRQ and the Effective Health Care Program. J Clin Epidemiol. 2011 Nov;64(11):1198-207. PMID: 21463926.

19. Chang SA, Choi SI, Choi EK, et al. Usefulness of 64-slice multidetector computed tomography as an initial diagnostic approach in patients with acute chest pain. Am Heart J. 2008 Aug;156(2):375-83. PMID: 18657674.

20. Sabharwal NK, Stoykova B, Taneja AK, et al. A randomized trial of exercise treadmill ECG versus stress SPECT myocardial perfusion imaging as an initial diagnostic strategy in stable patients with chest pain and suspected CAD: cost analysis. [Erratum appears in J Nucl Cardiol. 2007 May-Jun;14(3):414]. J Nucl Cardiol. 2007 Apr;14(2):174-86. PMID: 17386379.

21. Hoffmann U, Truong QA, Schoenfeld DA, et al. Coronary CT angiography versus standard evaluation in acute chest pain. N Engl J Med. 2012 Jul 26;367(4):299-308. PMID: 22830462.

22. Truong QA, Hayden D, Woodard PK, et al. Sex differences in the effectiveness of early coronary computed tomographic angiography compared with standard emergency department evaluation for acute chest pain: the Rule-Out Myocardial Infarction with Computer-Assisted Tomography (ROMICAT)-II Trial. Circulation. 2013 Jun 25;127(25):2494-502. PMID: 23685743.

23. Gruettner J, Fink C, Walter T, et al. Coronary computed tomography and triple rule out CT in patients with acute chest pain and an intermediate cardiac risk profile. Part 1: impact on patient management. Eur J Radiol. 2013 Jan;82(1):100-5. PMID: 22749769.

24. Henzler T, Gruettner J, Meyer M, et al. Coronary computed tomography and triple rule out CT in patients with acute chest pain and an intermediate cardiac risk for acute coronary syndrome: part 2: economic aspects. Eur J Radiol. 2013 Jan;82(1):106-11. PMID: 22835878.

25. Douglas PS, Hoffmann U, Patel MR, et al. Outcomes of anatomical versus functional testing for coronary artery disease. N Engl J Med. 2015 Apr 2;372(14):1291-300. PMID: 25773919.

26. Levsky JM, Spevack DM, Travin MI, et al. Coronary computed tomography angiography versus radionuclide myocardial perfusion imaging in patients with chest pain admitted to telemetry: a randomized trial. Ann Intern Med. 2015 Aug 4;163(3):174-83. PMID: 26052677.

27. Shaw LJ, Mieres JH, Hendel RH, et al. Comparative effectiveness of exercise electrocardiography with or without myocardial perfusion single photon emission computed tomography in women with suspected coronary artery disease: results from the What Is the Optimal Method for Ischemia Evaluation in Women (WOMEN) trial. Circulation. 2011 Sep 13;124(11):1239-49. PMID: 21844080.

28. Litt HI, Gatsonis C, Snyder B, et al. CT angiography for safe discharge of patients with possible acute coronary syndromes. N Engl J Med. 2012 Apr 12;366(15):1393-403. PMID: 22449295.

29. Miller AH, Pepe PE, Peshock R, et al. Is coronary computed tomography angiography a resource sparing strategy in the risk stratification and evaluation of acute chest pain? Results of a randomized controlled trial. Acad Emerg Med. 2011 May;18(5):458-67. PMID: 21569165.

30. Poon M, Cortegiano M, Abramowicz AJ, et al. Associations between routine coronary computed tomographic angiography and reduced unnecessary hospital admissions, length of stay, recidivism rates, and invasive coronary angiography in the emergency department triage of chest pain. J Am Coll Cardiol. 2013 Aug 6;62(6):543-52. PMID: 23684682.

31. Goldstein JA, Chinnaiyan KM, Abidov A, et al. The CT-STAT (Coronary Computed Tomographic Angiography for Systematic Triage of Acute Chest Pain Patients to Treatment) trial. J Am Coll Cardiol. 2011 Sep 27;58(14):1414-22. PMID: 21939822.

32. Goldstein JA, Gallagher MJ, O'Neill WW, et al. A randomized controlled trial of multi-slice coronary computed tomography f or evaluation of acute chest pain. J Am Coll Cardiol. 2007 Feb 27;49(8):863-71. PMID: 17320744.

33. Cheezum MK, Hulten EA, Taylor AJ, et al. Cardiac CT angiography compared with myocardial perfusion stress testing on downstream resource utilization. J Cardiovasc Comput Tomogr. 2011 Mar-Apr;5(2):101-9. PMID: 21256102.

34. Hamilton-Craig C, Fifoot A, Hansen M, et al. Diagnostic performance and cost of CT angiography versus stress ECG--a randomized prospective study of suspected acute coronary syndrome chest pain in the emergency department (CT-COMPARE). Int J Cardiol. 2014 Dec 20;177(3):867-73. PMID: 25466568.

35. Nielsen LH, Markenvard J, Jensen JM, et al. Frontline diagnostic evaluation of patients suspected of angina by coronary computed tomography reduces downstream resource utilization when compared to conventional ischemia testing. Int J Cardiovasc Imaging. 2011 Jul;27(6):813-23. PMID: 21042860.

36. Nielsen LH, Olsen J, Markenvard J, et al. Effects on costs of frontline diagnostic evaluation in patients suspected of angina: coronary computed tomography angiography vs. conventional ischaemia testing. Eur Heart J Cardiovasc Imaging. 2013 May;14(5):449-55. PMID: 22922828.

37. Hachamovitch R, Nutter B, Hlatky MA, et al. Patient management after noninvasive cardiac imaging results from SPARC (Study of Myocardial Perfusion and Coronary Anatomy Imaging Roles in Coronary Artery Disease). J Am Coll Cardiol. 2012 Jan 31;59(5):462-74. PMID: 22281249.

38. Hlatky MA, Shilane D, Hachamovitch R, et al. Economic outcomes in the Study of Myocardial Perfusion and Coronary Anatomy Imaging Roles in Coronary Artery Disease registry: the SPARC Study. J Am Coll Cardiol. 2014 Mar 18;63(10):1002-8. PMID: 24636556.

39. Min JK, Koduru S, Dunning AM, et al. Coronary CT angiography versus myocardial perfusion imaging for near-term quality of life, cost and radiation exposure: a prospective multicenter randomized pilot trial. J Cardiovasc Comput Tomogr. 2012 Jul-Aug;6(4):274-83. PMID: 22732201.

40. McKavanagh P, Lusk L, Ball PA, et al. A comparison of cardiac computerized tomography and exercise stress electrocardiogram test for the investigation of stable chest pain: the clinical results of the CAPP randomized prospective trial. Eur Heart J Cardiovasc Imaging. 2015 Apr; 6(4):441-8. PMID: 25473041.

41. Shreibati JB, Baker LC, Hlatky MA. Association of coronary CT angiography or stress testing with subsequent utilization and spending among Medicare beneficiaries. JAMA. 2011 Nov 16;306(19):2128-36. PMID: 22089720.

42. Tandon V, Hall D, Yam Y, et al. Rates of downstream invasive coronary angiography and revascularization: computed tomographic coronary angiography vs. Tc-99m single photon emission computed tomography. Eur Heart J. 2012 Mar;33(6):776-82. PMID: 21893487.

43. Min JK, Shaw LJ, Berman DS, et al. Costs and clinical outcomes in individuals without known coronary artery disease undergoing coronary computed tomographic angiography from an analysis of Medicare category III transaction codes. Am J Cardiol. 2008 Sep 15;102(6):672-8. PMID: 18773986.

44. Yamauchi T, Tamaki N, Kasanuki H, et al. Optimal initial diagnostic strategies for the evaluation of stable angina patients: a multicenter, prospective study on myocardial perfusion imaging, computed tomographic angiography, and coronary angiography. Circ J. 2012;76(12):2832-9. PMID: 22975716.

45. Sanfilippo AJ, Abdollah H, Knott TC, et al. Stress echocardiography in the evaluation of women presenting with chest pain syndrome: a randomized, prospective comparison with electrocardiographic stress testing. Can J Cardiol. 2005 Apr;21(5):405-12. PMID: 15861257.

46. Marwick TH, Shaw L, Case C, et al. Clinical and economic impact of exercise electrocardiography and exercise echocardiography in clinical practice. Eur Heart J. 2003 Jun;24(12):1153-63. PMID: 12804930.

47. Ollendorf DA, Colby JA, Cameron C, et al. Cardiac Nuclear Imaging. Final Appraisal Document. Prepared by the Institute for Clinical and Economic Review. Olympia, WA: Health Technology Assessment Program, Washington State Health Care Authority; August 12, 2013. www.hca.wa.gov/hta/documents/cni_final_report_081313.pdf.

48. Ollendorf DA, Göhler A, Pearson SD. Coronary Computed Tomographic Angiography for Detection of Coronary Artery Disease. Draft Appraisal Document. Institute for Clinical and Economic Review; October 17 2008. www.icer-review.org/wp-content/uploads/2013/04/CCTA_Rpt_Draft.pdf.

49. Dolor RJ, Patel MR, Melloni C, et al. 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. www.effectivehealthcare.ahrq.gov/ reports/final.cfm.

50. Mujtaba S, Pena JM, Pamerla M, et al. Publication trends in noninvasive cardiovascular imaging: 1991-2011: a retrospective observational study. Am J Cardiovasc Dis. 2013;3(4):247-54. PMID: 24224136.

51. Ronan G, Wolk MJ, Bailey SR, et al. ACCF/AHA/ASE/ASNC/HFSA/HRS/SCAI/SCCT/SCMR/STS 2013 Multimodality Appropriate Use Criteria for the Detection and Risk Assessment of Stable Ischemic Heart Disease: a report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Heart Association, American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Failure Society of America, Heart Rhythm Society, Society for Cardiovascular Angiography and Interventions, Society of Cardiovascular Computed Tomography, Society for Cardiovascular Magnetic Resonance, and Society of Thoracic Surgeons. J Nucl Cardiol. 2014;21(1):192-220. PMID: 24374980.

52. American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Society of Echocardiography, American Heart Association, et al. ACCF/ASE/AHA/ASNC/HFSA/HRS/SCAI/SCCM/SCCT/SCMR 2011 Appropriate Use Criteria for Echocardiography. A report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, American Society of Echocardiography, American Heart Association, American Society of Nuclear Cardiology, Heart Failure Society of America, Heart Rhythm Society, Society for Cardiovascular Angiography and Interventions, Society of Critical Care Medicine, Society of Cardiovascular Computed Tomography, Society for Cardiovascular Magnetic Resonance, American College of Chest Physicians. J Am Soc Echocardiogr. 2 011 Mar;24(3):229-67. PMID: 21338862.

53. Gibbons RJ, Balady GJ, Bricker JT, et al. ACC/AHA 2002 guideline update for exercise testing: summary article: a report of the American College of Cardiology/ American Heart Association Task Force on Practice Guidelines (Committee to Update the 1997 Exercise Testing Guidelines). Circulation. 2002 Oct;106(14):1883-92. PMID: 12356646.

54. Patel MR, Bailey SR, Bonow RO, et al. ACCF/SCAI/AATS/AHA/ASE/ASNC/HFSA/HRS/SCCM/SCCT/SCMR/STS 2012 Appropriate Use Criteria for Diagnostic Catheterization: a report of the American College of Cardiology Foundation Appropriate Use Criteria Task Force, Society for Cardiovascular Angiography and Interventions, American Association for Thoracic Surgery, American Heart Association, American Society of Echocardiography, American Society of Nuclear Cardiology, Heart Failure Society of America, Heart Rhythm Society, Society of Critical Care Medicine, Society of Cardiovascular Computed Tomography, Society for Cardiovascular Magnetic Resonance, and Society of Thoracic Surgeons. J Am Coll Cardiol. 2012 May 29;59(22):1995-2027. PMID: 22578925.

55. Pellikka PA, Nagueh SF, Elhendy AA, et al. American Society of Echocardiography recommendations for performance, interpretation, and application of stress echocardiography. J Am Soc Echocardiogr. 2007 Sep;20(9):1021-41. PMID: 17765820.

56. Shaw LJ, Berman DS, Picard MH, et al. Comparative definitions for moderate-severe ischemia in stress nuclear, echocardiography, and magnetic resonance imaging. JACC Cardiovasc Imaging. 2014 Jun;7(6):593-604. PMID: 24925328.

57. 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.

58. Rauch G, Rauch B, Schuler S, et al. Opportunities and challenges of clinical trials in cardiology using composite primary endpoints. World J Cardiol. 2015 Jan 26;7(1):1-5. PMID: 25632312.

59. Wasfy MM, Brady TJ, Abbara S, et al. Comparison of the Diamond-Forrester method and Duke Clinical Score to predict obstructive coronary artery disease by computed tomographic angiography. Am J Cardiol. 2012 Apr 1;109(7):998-1004. PMID: 22236462.

Full Report

This executive summary is part of the following document: Skelly AC, Hashimoto R, Buckley DI, Brodt ED, Noelck N, Totten AM, Lindner JR, Fu R, McDonagh M. Noninvasive Testing for Coronary Artery Disease. Comparative Effectiveness Review No. 171. (Prepared by the Pacific Northwest Evidence-based Practice Center under Contract No. 290-2012-00014-I.) AHRQ Publication No. 16-EHC011-EF. Rockville, MD: Agency for Healthcare Research and Quality; March 2016. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

AHRQ Pub. No. 16-EHC011-1-EF
March 2016