

NATIONAL QUALITY FORUM
National Voluntary Consensus Standards for Imaging Efficiency
Measure Summary

Measure Number: IEP-016-10

Measure Title: Cardiac stress imaging not meeting appropriate use criteria: Testing in asymptomatic, low risk patients

Description: Percentage of all stress SPECT MPI, stress echo, CCTA, and CMR performed in asymptomatic, low CHD risk patients for initial detection and risk assessment

Numerator Statement: Number of stress SPECT MPI, stress echo, CCTA, and CMR performed for asymptomatic, low CHD risk patients for initial detection and risk assessment*

Denominator statement: Number of stress SPECT MPI, stress echo, CCTA, and CMR performed

Level of Analysis: Facility/Agency

Data Source: Paper medical record/flowsheet, Survey: Patient

Measure developer: American College of Cardiology Foundation

Type of Endorsement (full or time-limited): Full Endorsement

Attachments: Imaging Efficiency Measures Micro-specifications 121809-633972762622979102 and Optimization of Patient Selection for Cardiac Stress Imaging-633972762344076422

NATIONAL QUALITY FORUM

Measure Evaluation 4.1 January 2010

This form contains the measure information submitted by stewards. Blank fields indicate no information was provided. Attachments also may have been submitted and are provided to reviewers. The sub-criteria and most of the footnotes from the [evaluation criteria](#) are provided in Word comments and will appear if your cursor is over the highlighted area (or in the margin if your Word program is set to show revisions in balloons). Hyperlinks to the evaluation criteria and ratings are provided in each section.

TAP/Workgroup (if utilized): Complete all **yellow highlighted** areas of the form. Evaluate the extent to which each sub-criterion is met. Based on your evaluation, summarize the strengths and weaknesses in each section.

Note: *If there is no TAP or workgroup, the SC also evaluates the sub-criteria (yellow highlighted areas).*

Steering Committee: Complete all **pink** highlighted areas of the form. Review the workgroup/TAP assessment of the sub-criterion, noting any areas of disagreement; then evaluate the extent to which each major criterion is met; and finally, indicate your recommendation for the endorsement. Provide the rationale for your ratings.

Evaluation ratings of the extent to which the criteria are met

- C = Completely (unquestionably demonstrated to meet the criterion)
- P = Partially (demonstrated to partially meet the criterion)
- M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
- N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
- NA = Not applicable (only an option for a few sub-criteria as indicated)

(for NQF staff use) NQF Review #: IEP-016-10 NQF Project: Efficiency: Imaging Efficiency	
MEASURE DESCRIPTIVE INFORMATION	
De.1 Measure Title: Cardiac stress imaging not meeting appropriate use criteria: Testing in asymptomatic, low risk patients	
De.2 Brief description of measure: Percentage of all stress SPECT MPI, stress echo, CCTA, and CMR performed in asymptomatic, low CHD risk patients for initial detection and risk assessment	
1.1-2 Type of Measure: efficiency/cost	
De.3 If included in a composite or paired with another measure, please identify composite or paired measure	
De.4 National Priority Partners Priority Area: Overuse	
De.5 IOM Quality Domain: efficiency	
De.6 Consumer Care Need: Living With Illness	

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	NQF Staff
<p>A. The measure is in the public domain or an intellectual property (measure steward agreement) is signed. <i>Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.</i></p> <p>A.1 Do you attest that the measure steward holds intellectual property rights to the measure and the right to use aspects of the measure owned by another entity (e.g., risk model, code set)? Yes</p> <p>A.2 Indicate if Proprietary Measure (as defined in measure steward agreement): proprietary measure</p> <p>A.3 Measure Steward Agreement: agreement signed and submitted</p> <p>A.4 Measure Steward Agreement attached: MSA_ACCF-633983115368409334.pdf</p>	<p>A</p> <p>Y <input type="checkbox"/></p> <p>N <input type="checkbox"/></p>

B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. Yes, information provided in contact section	B Y <input type="checkbox"/> N <input type="checkbox"/>
C. The intended use of the measure includes <u>both</u> public reporting <u>and</u> quality improvement. ► Purpose: public reporting, quality improvement Accreditation	C Y <input type="checkbox"/> N <input type="checkbox"/>
D. The requested measure submission information is complete. Generally, measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. Measures that have not been tested are only potentially eligible for a time-limited endorsement and in that case, measure owners must verify that testing will be completed within 12 months of endorsement. D.1 Testing: Yes, fully developed and tested D.2 Have NQF-endorsed measures been reviewed to identify if there are similar or related measures? Yes	D Y <input type="checkbox"/> N <input type="checkbox"/>
(for NQF staff use) Have all conditions for consideration been met? Staff Notes to Steward (if submission returned):	Met Y <input type="checkbox"/> N <input type="checkbox"/>
Staff Notes to Reviewers (issues or questions regarding any criteria):	
Staff Reviewer Name(s):	

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. <i>Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria. (evaluation criteria)</i> 1a. High Impact	Eval Rating
(for NQF staff use) Specific NPP goal:	
1a.1 Demonstrated High Impact Aspect of Healthcare: a leading cause of morbidity/mortality, high resource use, frequently performed procedure, patient/societal consequences of poor quality 1a.2 1a.3 Summary of Evidence of High Impact: Cardiac imaging is a mainstay in medical decision-making for patients with known or suspected heart disease. However, expenditures related to imaging comprise a significant portion of the health care budget. Much scrutiny has been focused on cardiovascular imaging with regard to the potential for overuse, especially in view of substantial geographic variation in ordering patterns and the limited amount of evidence-based data supporting the use of imaging as it relates to patient outcomes. Given the significant contribution of heart disease to morbidity and mortality and the prevalence of cardiovascular disease, it is important to determine the appropriate use of diagnostic tests such as stress echocardiography, stress SPECT MPI, CCTA, and CMR . 1a.4 Citations for Evidence of High Impact: Patel MR, Spertus JA, Brindis RG., et al. "ACCF proposed method for evaluating the appropriateness of cardiovascular imaging." J Am Coll Cardiol. 2005 Oct 18;46(8):1606-13. Hendel RC. The revolution and evolution of appropriateness in cardiac imaging J Nucl Cardiol 2008;15:494-496. Hendel RC. Utilization management of cardiovascular imaging p-certification and appropriateness J Am Coll	1a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>

Comment [KP1]: 1a. The measure focus addresses:
 • a specific national health goal/priority identified by NQF's National Priorities Partners; OR
 • a demonstrated high impact aspect of healthcare (e.g., affects large numbers, leading cause of morbidity/mortality, high resource use (current and/or future), severity of illness, and patient/societal consequences of poor quality).

<p>Cardiol Img 2008;1:241-248.</p> <p>Iglehart JK. Health insurers and medical-imaging policy—a work in progress N Engl J Med 2009;360:1030-1037.</p> <p>Iglehart JK. The new era of medical imaging—progress and pitfalls N Engl J Med 2006;354:2822-2828</p> <p>Krumholz HM, Keenan PS, Brush JE et al. Standards for measures used for public reporting of efficiency in health care. J Am Coll Cardiol. 2008 Oct 28;52(18):1518-26.</p> <p>Hendel, RC; Cerqueira, M; Douglas, PS et al. "A Multicenter Assessment of the Use of Single-Photon Emission Computed Tomography Myocardial Perfusion Imaging With Appropriateness Criteria". J Am Coll Cardiol. Published online December 10, 2009.</p> <p>Mehta R, Agarwal S, Chandra S, Ward RP, Williams KA: Evaluation of the American College of Cardiology Foundation/American Society of Nuclear Cardiology appropriateness criteria for SPECT myocardial perfusion imaging. J Nucl Cardiol. 2008;5:337-44.</p> <p>Ward RP, Al-Mallah MH, Grossman GB, Hansen CL, Hendel RC, Kerwin TC, McCallister BD Jr., Mehta R, Dm Polk, Tilkemeier PL, Vashist A, Williams KA, Wolinsky DG, Ficaro EP: American Society of Nuclear Cardiology: American Society of Nuclear Cardiology review of the ACCF/ASNC appropriateness criteria for single-photon emission computed tomography myocardial perfusion imaging (SPECT MPI). J Nucl Cardiol. 2007;14:e26-38.</p> <p>Gibbons RJ, Miller TD, Hodge D, Urban L, Araoz PA, Pellikka P, McCully RB: Application of appropriateness criteria to stress single photon emission computed tomography sestamibi studies and stress echocardiograms in an academic medical center. J Am Coll Cardiol. 2008;51:1283-9.</p> <p>McCully RB, Pellikka PA, Hodge DO, et al: Applicability of Appropriateness Criteria for Stress Imaging. Circ Cardiovasc Imaging. 2009;2:213-8.</p>	
<p>1b. Opportunity for Improvement</p> <p>1b.1 Benefits (improvements in quality) envisioned by use of this measure: Appropriate use criteria define “when to do” and “how often to do” a given procedure in the context of scientific evidence, the health care environment, the patient’s profile and a physician’s judgment. While practice guidelines provide a foundation for summarizing evidence-based cardiovascular care or for providing expert consensus opinions, in many areas, marked variability remains in the use of cardiovascular procedures, raising questions about over-use and under-use. Appropriate use criteria provide practical tools to measure this variability and to look at utilization patterns. The criteria are designed to examine the use of diagnostic and therapeutic procedures to support efficient use of medical resources, while also providing patients with quality, appropriate care.</p> <p>A measure that reports rates of inappropriate imaging within practices would contain information regarding both cost and quality, because an inappropriate test results in both higher costs and poorer-quality care. Conversely, a reduction in this rate would simultaneously improve quality and decrease cost. Improvements in this metric should lead to consistent application of AUC and improve the efficiency of the system.</p> <p>1b.2 Summary of data demonstrating performance gap (variation or overall poor performance) across providers: A number of studies have indicated that cardiac imaging in low CHD risk, asymptomatic patients is a common inappropriate use</p> <p>1b.3 Citations for data on performance gap: Krumholz HM, Keenan PS, Brush JE et al. Standards for measures used for public reporting of efficiency in health care. J Am Coll Cardiol. 2008 Oct 28;52(18):1518-26.</p> <p>Hendel, RC; Cerqueira, M; Douglas, PS et al. "A Multicenter Assessment of the Use of Single-Photon Emission Computed Tomography Myocardial Perfusion Imaging With Appropriateness Criteria". J Am Coll</p>	<p style="text-align: center;">1b C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>

Comment [KP2]: 1b. Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating considerable variation, or overall poor performance, in the quality of care across providers and/or population groups (disparities in care).

Comment [k3]: 1 Examples of data on opportunity for improvement include, but are not limited to: prior studies, epidemiologic data, measure data from pilot testing or implementation. If data are not available, the measure focus is systematically assessed (e.g., expert panel rating) and judged to be a quality problem.

Cardiol. Published online December 10, 2009.

Mehra R, Agarwal S, Chandra S, Ward RP, Williams KA: Evaluation of the American College of Cardiology Foundation/American Society of Nuclear Cardiology appropriateness criteria for SPECT myocardial perfusion imaging. J Nucl Cardiol. 2008;5:337-44.

Ward RP, Al-Mallah MH, Grossman GB, Hansen CL, Hendel RC, Kerwin TC, McCallister BD Jr., Mehta R, Dm Polk, Tilkemeier PL, Vashist A, Williams KA, Wolinsky DG, Ficaro EP: American Society of Nuclear Cardiology: American Society of Nuclear Cardiology review of the ACCF/ASNC appropriateness criteria for single-photon emission computed tomography myocardial perfusion imaging SPECT MPI). J Nucl Cardiol. 2007;14:e26-38.

Gibbons RJ, Miller TD, Hodge D, Urban L, Araoz PA, Pellikka P, McCully RB: Application of appropriateness criteria to stress single photon emission computed tomography sestamibi studies and stress echocardiograms in an academic medical center. J Am Coll Cardiol. 2008;51:1283-9.

McCully RB, Pellikka PA, Hodge DO, et al: Applicability of Appropriateness Criteria for Stress Imaging. Circ Cardiovasc Imaging. 2009;2:213-8.

1b.4 Summary of Data on disparities by population group:
None

1b.5 Citations for data on Disparities:
None

1c. Outcome or Evidence to Support Measure Focus

1c.1 Relationship to Outcomes (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population*): Diagnostic testing, such as stress SPECT MPI, stress echocardiography, CCTA, and CMR, is used to detect disease and provide risk assessment used to modify treatment strategies and approaches. Information provided by such testing can initiate, modify and stop further treatments for coronary heart disease (medications and revascularization) which have an impact on patient outcomes. In addition, false positives and false negatives can adversely impact the patient and their treatment outcomes. Lastly, radiation from stress SPECT MPI poses a minimal but still important consideration for patient safety. Ensuring proper patient selection can avoid using resources in patients not expected to benefit from the testings and for which the associated risks would be unnecessary.

1c.2-3. Type of Evidence: evidence based guideline, expert opinion

1c.4 Summary of Evidence (*as described in the criteria; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*):

Appropriate use criteria define "when to do" and "how often to do" a given procedure in the context of scientific evidence, the health care environment, the patient's profile and a physician's judgment. While practice guidelines provide a foundation for summarizing evidence-based cardiovascular care or for providing expert consensus opinions, in many areas, marked variability remains in the use of cardiovascular procedures, raising questions about over-use and under-use. Appropriate use criteria provide practical tools to measure this variability and to look at utilization patterns. The criteria are designed to examine the use of diagnostic and therapeutic procedures to support efficient use of medical resources, while also providing patients with quality, appropriate care.

Because of its patient-centered approach, it is hoped that appropriate use criteria can lead to patient education regarding expected benefits and risks associated with diagnostic and therapeutic procedures. In addition, physicians, payers and medical facilities can use the criteria prospectively or retrospectively to assess practice patterns, design ordering protocols and/or provide the basis for quality improvement activities focused on ensuring the most appropriate care for patients.

Unlike many performance measures which have primarily focused on underuse of evidence based therapies, this measure set focuses on overuse of diagnostic technology for which there are a limited number of

Comment [k4]: 1c. The measure focus is:
 •an outcome (e.g., morbidity, mortality, function, health-related quality of life) that is relevant to, or associated with, a national health goal/priority, the condition, population, and/or care being addressed;
 OR
 •if an intermediate outcome, process, structure, etc., there is evidence that supports the specific measure focus as follows:
 oIntermediate outcome - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
 oProcess - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and
 if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
 oStructure - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
 oPatient experience - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
 oAccess - evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
 oEfficiency - demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

Comment [k5]: 4 Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multi-step process, the step with the greatest effect on the desired outcome should be selected as the focus of measurement. For example, although assessment of immunization status and recommending immunization are necessary steps, they are not sufficient to achieve the desired impact on health status - patients must be vaccinated to achieve immunity. This does not preclude consideration of measures of preventive screening interventions where there is a strong link with desired outcomes (e.g., mammography) or measures for multiple care processes that affect a single outcome.

1c
 C
 P
 M
 N

prospective, randomized trials. As such, typical guideline based approaches to selecting the measures such as focusing on Class I, Level of Evidence A recommendations is not feasible. However, appropriate use criteria were designed to highlight patient scenarios for which observational data and expert opinion would indicate the incremental benefit gained by use of a diagnostic test is not justified. The data supporting these common inappropriate indications is based on well known risk algorithms (with more than 30 years of use) and observational data describing retrospectively and prospectively how the criteria have performed in determining test yield characteristics. In both cases, the evidence indicates that the expected information gained from a diagnostic test would be minimal compared to other patient populations. As such, the expected mortality benefit and treatment impact that such diagnostic testing would have based on current treatment guidelines would be minimal, as well. Table 1 highlights the yield of testing in a recent prospective study applying the SPECT MPI Appropriate Use Criteria. While the overall test yield was low in this study, a test with no or small amounts of ischemia in certain populations can be very meaningful and impact on the decision of whether to avoid more aggressive medical or interventional therapies. However, in the patient populations highlighted by this measure set, it would be difficult to justify the value of a negative test as these patients already would be unlikely to be candidates for therapy. Furthermore, a positive test, as can be seen below, is very unlikely and is an even more infrequent occurrence than in other patient scenarios.

Table 1. Test Results (All Sites, All Dates); N= 6333
Stratified by Appropriate Use Classification (ACC Internal Data Analysis of pilot data from: Hendel RC, Cerqueira M, Douglas PS, et al. A Multicenter Assessment of the Use of Single-Photon Emission Computed Tomography Myocardial Perfusion Imaging With Appropriateness Criteria. JACC. December 10, 2009.)

Top Inappropriate
(low risk – Asx; routine post PCI; periop eval for low risk surgery)

Moderate Ischemia
0.6% of all patients in pilot N=38

Severe Ischemia
0.1% of all patients in pilot N=7

Moderate or Severe Ischemia
0.7% of all patients in pilot N=45

All other patients in pilot
Moderate Ischemia
6.1% of all patients in pilot N=384

Severe Ischemia
1.0% of all patients in pilot N=62

Moderate or Severe Ischemia
7.1% of all patients in pilot N=446

Note: No perioperative low risk surgery patients had moderate to severe ischemia

1c.5 Rating of strength/quality of evidence (also provide narrative description of the rating and by whom):

Observational Studies and Expert Opinion - ACCF AUC Criteria Task Force

1c.6 Method for rating evidence: Specific evidence grades are not assigned by Appropriate Use Criteria (AUC), but generally diagnostic imaging evidence is based on observational studies, including well known risk models such as Framingham and Diamond and Forrester. In addition, a RAND modified Delphi process is used to determine the AUC rating that combines expert opinion with available evidence and specific patient information. Few studies are conducted to demonstrate a lack of benefit and thus, clinical risk and

Comment [k6]: 3 The strength of the body of evidence for the specific measure focus should be systematically assessed and rated (e.g., USPSTF grading system <http://www.ahrq.gov/clinic/uspstf07/methods/benefit.htm>). If the USPSTF grading system was not used, the grading system is explained including how it relates to the USPSTF grades or why it does not. However, evidence is not limited to quantitative studies and the best type of evidence depends upon the question being studied (e.g., randomized controlled trials appropriate for studying drug efficacy are not well suited for complex system changes). When qualitative studies are used, appropriate qualitative research criteria are used to judge the strength of the evidence.

expert opinion is required to develop the AUC.

1c.7 Summary of Controversy/Contradictory Evidence: Risk algorithms are under regular review and scrutiny as to whether they properly reflect the patient populations being considered.

1c.8 Citations for Evidence (other than guidelines): Grundy SM, Pasternak R, Greenland P, et al. AHA/ACC scientific statement: assessment of cardiovascular risk by use of multiple-risk-factor assessment equations: a statement for healthcare professionals from the American Heart Association and the American College of Cardiology *J Am Coll Cardiol* 1999;34:1348-1359.

1c.9 Quote the Specific guideline recommendation (including guideline number and/or page number):

2002 Stable Angina Guideline

"Asymptomatic patients with abnormal findings on ambulatory ECG or EBCT who are able to exercise can be evaluated with exercise ECG testing, although the efficacy of exercise ECG testing in asymptomatic patients is not well established. Stress imaging procedures (i.e., either stress myocardial perfusion imaging or stress echocardiography) are generally not indicated in most such patients."

AUC Indications

2008 Appropriateness Criteria for Stress Echocardiography

Indication 11: Detection of CAD and Risk Assessment: Asymptomatic (without Chest Pain Syndrome or Anginal Equivalent): Low CHD risk (Framingham risk criteria) - Inappropriate (1)

2009 Appropriate Use Criteria for Cardiac Radionuclide Imaging

Indication 12: Detection of CAD/Risk Assessment Without Ischemic Equivalent: Asymptomatic: Low CHD risk (ATP III risk criteria) - Inappropriate (1)

2006 Appropriateness Criteria for CCT and CMR

Indication 10 - Detection of CAD: Asymptomatic (Use of CCTA) (Without Chest Pain Syndrome): Asymptomatic: Low CHD risk (Framingham risk criteria) - Inappropriate (1)

2002 Chronic Stable Angina Guideline

Class III

Recommendations for Cardiac Stress Imaging as the Initial Test for Diagnosis in Asymptomatic Patients

1. Exercise or dobutamine echocardiography in asymptomatic patients with left bundle-branch block. (Level of Evidence: C)
2. Exercise myocardial perfusion imaging, exercise echocardiography, adenosine or dipyridamole myocardial perfusion imaging, or dobutamine echocardiography as the initial stress test in an asymptomatic patient with a normal rest ECG who is not taking digoxin. (Level of Evidence: C)
3. Adenosine or dipyridamole myocardial perfusion imaging or dobutamine echocardiography in asymptomatic patients who are able to exercise and do not have left bundle-branch block or electronically paced ventricular rhythm. (Level of Evidence: C)

1c.10 Clinical Practice Guideline Citation: Gibbons RJ, Abrams J, Chatterjee K, et al. ACC/AHA 2002 guideline update for the management of patients with chronic stable angina—summary article: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee on the Management of Patients With Chronic Stable Angina) *J Am Coll Cardiol* 2003;41:159-168.

Douglas PS, Khandheria B, Stainback RF, ACCF/ASE/ACEP/AHA/ASNC/SCAI/SCCT/SCMR2008 appropriateness criteria for stress echocardiography. *J Am Coll Cardiol*. 2008;51:1127-47.

Hendel RH, Berman DS, Di Carli MF, et al. ACCF/ASNC/ACR/AHA/ASE/SCCT/SCMR/SNM 2009 Appropriate Use Criteria for Cardiac Radionuclide Imaging. *J Am Coll Cardiol*. 2009;53:2201-29.

Hendel RC, Patel MR, Kramer CM, Poon M. ACCF/ACR/SCCT/SCMR/ASNC/NASCI/SCAI/SIR 2006 appropriateness criteria for cardiac computed tomography and cardiac magnetic resonance imaging. *J Am*

<p>Coll Cardiol 2006;48:1475-97. 1c.11 National Guideline Clearinghouse or other URL: http://www.acc.org/qualityandscience/clinical/statements.htm</p> <p>1c.12 Rating of strength of recommendation <i>(also provide narrative description of the rating and by whom):</i> Class III - Conditions for which there is evidence and/or general agreement that the procedure/treatment is not useful/effective and in some cases may be harmful. AND Inappropriate - test is not generally acceptable and is not a reasonable approach for the indication</p> <p>1c.13 Method for rating strength of recommendation <i>(If different from USPSTF system, also describe rating and how it relates to USPSTF):</i> Guidelines - consensus development Appropriate Use Criteria - modified RAND Delphi Method</p> <p>1c.14 Rationale for using this guideline over others: Appropriate Use Criteria developed by ACC in partnership with other societies provide more specific patient scenarios and directly address inappropriate use while other guidelines and appropriateness criteria documents do not offer this specificity nor do they generally address overuse/waste.</p>	
<p>TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for <i>Importance to Measure and Report</i>?</p>	<p>1</p>
<p>Steering Committee: Was the threshold criterion, <i>Importance to Measure and Report</i>, met? Rationale:</p>	<p>1 Y <input type="checkbox"/> N <input type="checkbox"/></p>
<p>2. SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES</p>	
<p>Extent to which the measure, <u>as specified</u>, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)</p>	<p>Eval Rating</p>
<p>2a. MEASURE SPECIFICATIONS</p>	
<p>S.1 Do you have a web page where current detailed measure specifications can be obtained? S.2 If yes, provide web page URL:</p>	
<p>2a. Precisely Specified</p> <p>2a.1 Numerator Statement <i>(Brief, text description of the numerator - what is being measured about the target population, e.g. target condition, event, or outcome):</i> Number of stress SPECT MPI, stress echo, CCTA, and CMR performed for asymptomatic, low CHD risk patients for initial detection and risk assessment*</p> <p>2a.2 Numerator Time Window <i>(The time period in which cases are eligible for inclusion in the numerator):</i> Sample of all SPECT MPI, stress echo, CCTA, and CMR test orders during a calendar year using a single, consecutive 60 day time period</p> <p>2a.3 Numerator Details <i>(All information required to collect/calculate the numerator, including all codes, logic, and definitions):</i> For all orders in asymptomatic patients, determine orders for initial diagnosis and risk assessment. In doing so, patients with known CHD, prior PCI or prior CABG and the following exclusions are not included.</p> <p>Patients qualify for this numerator if: - Asymptomatic AND - Low CHD risk based on clinician estimate AND NOT any of the following: - Known CAD, including • prior MI</p>	<p>2a- specs C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>

Comment [k7]: USPSTF grading system <http://www.ahrq.gov/clinic/uspstf/grades.htm>: A - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. B - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. C - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. D - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. I - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

Comment [KP8]: 2a. The measure is well defined and precisely specified so that it can be implemented consistently within and across organizations and allow for comparability. The required data elements are of high quality as defined by NOF's Health Information Technology Expert Panel (HITEP).

- prior ACS
- prior CABG
- prior PCI or
- CHD on prior diagnostic test
- Exercise stress treadmill
- Non-invasive imaging
- Stress echo
- Stress SPECT MPI
- CT Angiography
- Calcium Scoring
- Invasive imaging (cardiac catheterization)
- Ischemic equivalent
- Undergone prior CHD assessment by one the following methods no matter the test result:
 - o Exercise stress treadmill
 - o Non-invasive imaging
 - Stress echo
 - Stress SPECT MPI
 - CT Angiography
 - Calcium Scoring
 - o Invasive imaging (cardiac catheterization)
- Patients for whom preoperative testing is the primary reason for imaging

Submission of individual clinical data variables required for Framingham risk (ATP III criteria) calculation for asymptomatic patients is recognized to place a significant data collection burden upon institutions and may not be possible based on data elements that are readily available at the imaging laboratory. As such, a clinician estimate of CHD risk will be collected for all asymptomatic patients who are being seen for initial detection and risk assessment without known coronary heart disease. However, in making their estimate, clinicians should consider the maximum number of available patient factors used to estimate risk based on Framingham (ATP III criteria), typically age, gender, diabetes, smoking status, and use of blood pressure medication, and integrate age appropriate estimates for missing elements, such as LDL or standard blood pressure. While calculation of the estimate does not require submission of the actual clinical data elements other than the clinician estimate of CHD risk, clinicians are attesting to the accuracy of the estimate by submitting it. An audit of clinician estimates should be completed on a subset of clinicians to verify their estimates as being accurate based on the data that was available.

NOTE: Data collection from patient requisition is required to adequately determine patient's symptom status and clinical risk. Determination with only administrative data is not possible for this measure.

2a.4 Denominator Statement (*Brief, text description of the denominator - target population being measured*):

Number of stress SPECT MPI, stress echo, CCTA, and CMR performed

2a.5 Target population gender: Male, Female

2a.6 Target population age range: 18 years old and older - Appropriate Use Criteria only developed for adults

2a.7 Denominator Time Window (*The time period in which cases are eligible for inclusion in the denominator*):

Sample of all stress SPECT MPI, stress echo, CCTA, and CMR test orders during a calendar year using a single, consecutive 60 day time period

2a.8 Denominator Details (*All information required to collect/calculate the denominator - the target population being measured - including all codes, logic, and definitions*):

All consecutive stress SPECT MPI, stress echocardiography, CCTA, and CMR orders

Measurement Entity: Imaging laboratory prospectively measured on test requisition forms and/or patient charts

Level of Measurement/Analysis: Imaging laboratory*

<p>*Attribution for inappropriate use is shared between the ordering physician and imaging laboratory. In an ideal world, attribution to the ordering physician or institution, as well as the imaging laboratory, would be reflected in the reporting of these measures. However, there are numerous complexities that prevent assignment of these measures to individual ordering physicians. For example, ordering volumes from individual physicians and institutions are insufficient to make meaningful comparisons to allow such attribution. Thus, these measures will be reported at the level of the imaging laboratory. However, the extent to which the institution housing the imaging laboratory can impact these measures will be dependent upon cooperation of ordering physicians with the imaging laboratory.</p>
<p>2a.9 Denominator Exclusions <i>(Brief text description of exclusions from the target population):</i></p> <p>2a.10 Denominator Exclusion Details <i>(All information required to collect exclusions to the denominator, including all codes, logic, and definitions):</i></p>
<p>2a.11 Stratification Details/Variables <i>(All information required to stratify the measure including the stratification variables, all codes, logic, and definitions):</i> None.</p>
<p>2a.12-13 Risk Adjustment Type: no risk adjustment necessary</p> <p>2a.14 Risk Adjustment Methodology/Variables <i>(List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method):</i> N/A</p> <p>2a.15-17 Detailed risk model available Web page URL or attachment:</p>
<p>2a.18-19 Type of Score: rate/proportion 2a.20 Interpretation of Score: better quality = lower score 2a.21 Calculation Algorithm <i>(Describe the calculation of the measure as a flowchart or series of steps):</i> Locate all stress SPECT MPI, stress echocardiography, CCTA, and CMR orders performed during the sampling period. Record the total number of tests during the sampling period as the denominator. From this sets of test orders, identify orders containing the criteria listed in the numerator</p>
<p>2a.22 Describe the method for discriminating performance <i>(e.g., significance testing):</i> No laboratory is expected to achieve 0% inappropriate orders as there always will be extenuating circumstances not captured by the appropriate use criteria. However, it is expected that significance testing can be applied to differentiate performance between laboratories and for a given laboratory over time.</p>
<p>2a.23 Sampling (Survey) Methodology <i>If measure is based on a sample (or survey), provide instructions for obtaining the sample, conducting the survey and guidance on minimum sample size (response rate):</i> Measures are to be developed based on a sample of a full calendar year based on the following sampling methodology: Select a starting month: <input type="radio"/> January <input type="radio"/> March <input type="radio"/> May <input type="radio"/> July <input type="radio"/> September <input type="radio"/> November Begin 60 day data collection period on the 1st on the month for the selected starting month Determine whether at least 30 stress SPECT, stress echo, CCTA, or CMR orders have been placed during the</p>

Comment [k9]: 11 Risk factors that influence outcomes should not be specified as exclusions.
 12 Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

selected time period. If not, select another time period with a minimum number of 30 cases. If no time period includes the minimum number of cases, then the imaging laboratory does not have sufficient volume to report this measure.

Sampling is required for this measure as full year data collection does not alter performance rates for this measure and would place an additional data collection burden on laboratories. It also allows laboratories to share performance with ordering physicians more quickly than would be possible under full year calendar reporting.

2a.24 Data Source (Check the source(s) for which the measure is specified and tested)
 paper medical record/flowsheet, Survey: Patient

2a.25 Data source/data collection instrument (Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.):
 Optimization of Patient Selection for Cardiac Imaging

2a.26-28 Data source/data collection instrument reference web page URL or attachment: Attachment Optimization of Patient Selection for Cardiac Imaging-634094461412753144.doc

2a.29-31 Data dictionary/code table web page URL or attachment: Attachment Imaging Efficiency Measures Micro-specifications 121809-633972762622979102.doc

2a.32-35 Level of Measurement/Analysis (Check the level(s) for which the measure is specified and tested)
 Facility/Agency

2a.36-37 Care Settings (Check the setting(s) for which the measure is specified and tested)
 Ambulatory Care: Office, Ambulatory Care: Hospital Outpatient

2a.38-41 Clinical Services (Healthcare services being measured, check all that apply)
 Imaging

TESTING/ANALYSIS

2b. Reliability testing

2b.1 Data/sample (description of data/sample and size): No direct reliability testing of these measures has been undertaken. However, reliability testing has been performed on individual patient populations as a part of studies related to implementation of appropriate use criteria. In addition, a number of studies have demonstrated remarkable consistency in their findings related to the proportion of inappropriate studies and the relative frequency of common inappropriate indications.

2b.2 Analytic Method (type of reliability & rationale, method for testing):
 Agreement between 2 nurse abstracters - kappa test

Assignment of patients to appropriate use criteria indication requires the same data elements used to calculate this measure set. In fact, complete assignment of all patients and not just inappropriate patients requires a greater number of data variables for agreement between abstracters. It is anticipated that the more limited data set required for this measure set would yield even higher kappa values.

2b.3 Testing Results (reliability statistics, assessment of adequacy in the context of norms for the test conducted):

Gibbons RJ, Miller TD, Hodge D, Urban L, Araoz PA, Pellikka P, McCully RB: Application of appropriateness criteria to stress single photon emission computed tomography sestamibi studies and stress echocardiograms in an academic medical center. J Am Coll Cardiol. 2008;51:1283-9.

Nurse abstracter agreement kappa = .60 for indications other than repeat testing

2b
 C
 P
 M
 N

Comment [KP10]: 2b. Reliability testing demonstrates the measure results are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period.

Comment [k11]: 8 Examples of reliability testing include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing may address the data items or final measure score.

<p>McCully RB, Pellikka PA, Hodge DO, Araoz PA, Miller TD, Gibbons RJ. Applicability of appropriateness criteria for stress imaging: similarities and differences between stress echocardiography and single-photon emission computed tomography myocardial perfusion imaging criteria. <i>Circ Cardiovasc Imaging</i>. 2009 May;2(3):213-8. Nurse abstracter agreement kappa=0.72 for stress echocardiography</p>	
<p>2c. Validity testing</p> <p>2c.1 Data/sample (description of data/sample and size): 6,333 patients who underwent stress SPECT MPI testing as a part of the ACCF/UHC pilot</p> <p>2c.2 Analytic Method (type of validity & rationale, method for testing): Comparison of test yield as determined by extent of ischemia (moderate to severe ischemia - which would indicate a test finding that could potentially impact treatment) in the overall testing population:</p> <p>Patients Classified in the Top Inappropriate Indications Versus Patients not Classified in the Top Inappropriate Indications (Other)</p> <p>2c.3 Testing Results (statistical results, assessment of adequacy in the context of norms for the test conducted): Table 1. Test Results (All Sites, All Dates); N= 6333 Stratified by Appropriate Use Classification (ACC Internal Data Analysis of pilot data from: Hendel RC, Cerqueira M, Douglas PS, et al. A Multicenter Assessment of the Use of Single-Photon Emission Computed Tomography Myocardial Perfusion Imaging With Appropriateness Criteria. <i>JACC</i>. December 10, 2009.)</p> <p>Top Inappropriate (low risk - Asx; routine post PCI; periop eval for low risk surgery)</p> <p>Moderate Ischemia 0.6% of all patients in pilot N=38</p> <p>Severe Ischemia 0.1% of all patients in pilot N=7</p> <p>Moderate or Severe Ischemia 0.7% of all patients in pilot N=45</p> <p>All other patients in pilot Moderate Ischemia 6.1% of all patients in pilot N=384</p> <p>Severe Ischemia 1.0% of all patients in pilot N=62</p> <p>Moderate or Severe Ischemia 7.1% of all patients in pilot N=446</p> <p>Note: No perioperative low risk surgery patients had moderate to severe ischemia</p>	<p>2c C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/></p>
<p>2d. Exclusions Justified</p> <p>2d.1 Summary of Evidence supporting exclusion(s): Patients without sufficient patient selection criteria recorded in measure as described by: Adequacy of data to assess appropriate use of cardiac stress imaging</p>	<p>2d C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/></p>

Comment [KP12]: 2c. Validity testing demonstrates that the measure reflects the quality of care provided, adequately distinguishing good and poor quality. If face validity is the only validity addressed, it is systematically assessed.

Comment [k13]: 9 Examples of validity testing include, but are not limited to: determining if measure scores adequately distinguish between providers known to have good or poor quality assessed by another valid method; correlation of measure scores with another valid indicator of quality for the specific topic; ability of measure scores to predict scores on some other related valid measure; content validity for multi-item scales/tests. Face validity is a subjective assessment by experts of whether the measure reflects the quality of care (e.g., whether the proportion of patients with BP < 140/90 is a marker of quality). If face validity is the only validity addressed, it is systematically assessed (e.g., ratings by relevant stakeholders) and the measure is judged to represent quality care for the specific topic and that the measure focus is the most important aspect of quality for the specific topic.

Comment [KP14]: 2d. Clinically necessary measure exclusions are identified and must be:

- supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion;
- AND
- a clinically appropriate exception (e.g., contraindication) to eligibility for the measure focus;
- AND
- precisely defined and specified:
 - if there is substantial variability in exclusions across providers, the measure is specified so that exclusions are computable and the effect on the measure is transparent (i.e., impact clearly delineated, such as number of cases excluded, exclusion rates by type of exclusion);
 - if patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that it strongly impacts performance on the measure and the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

Comment [k15]: 10 Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, sensitivity analyses with and without the exclusion, and variability of exclusions across providers.

<p>A lack of data on patient selection criteria makes it impossible to determine whether a patient should have been counted as a part of numerator and thus, these patients also are not counted as a part of the denominator to ensure that these test orders do not dilute the patient population measured.</p> <p>2d.2 Citations for Evidence: None.</p> <p>2d.3 Data/sample (description of data/sample and size): None.</p> <p>2d.4 Analytic Method (type analysis & rationale): None.</p> <p>2d.5 Testing Results (e.g., frequency, variability, sensitivity analyses): None.</p>	
<p>2e. Risk Adjustment for Outcomes/ Resource Use Measures</p> <p>2e.1 Data/sample (description of data/sample and size): No risk adjustment</p> <p>2e.2 Analytic Method (type of risk adjustment, analysis, & rationale): No risk adjustment</p> <p>2e.3 Testing Results (risk model performance metrics): No risk adjustment</p> <p>2e.4 If outcome or resource use measure is not risk adjusted, provide rationale: No risk adjustment</p>	<p>2e</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p> <p>NA <input type="checkbox"/></p>
<p>2f. Identification of Meaningful Differences in Performance</p> <p>2f.1 Data/sample from Testing or Current Use (description of data/sample and size): 997 patients enrolled in ACC/UHC SPECT MPI AUC Pilot - 4 sites; patients enrolled in March and April (Internal ACC data analysis of Hendel, RC; Cerqueira, M; Douglas, PS et al. "A Multicenter Assessment of the Use of Single-Photon Emission Computed Tomography Myocardial Perfusion Imaging With Appropriateness Criteria". J Am Coll Cardiol. Published online December 10, 2009.)</p> <p>2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance (type of analysis & rationale): Data from all sites showed that this indication was among their most frequent inappropriate indications. The difference between the best and worst performers was more than twofold.</p> <p>2f.3 Provide Measure Scores from Testing or Current Use (description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance): Site 1 6.8% Site 2 8.8% Site 3 5.7% Site 4 3.5%</p>	<p>2f</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p>
<p>2g. Comparability of Multiple Data Sources/Methods</p> <p>2g.1 Data/sample (description of data/sample and size): None.</p> <p>2g.2 Analytic Method (type of analysis & rationale): None.</p> <p>2g.3 Testing Results (e.g., correlation statistics, comparison of rankings): It is anticipated that prospective data collection of some of these elements will be required using a physician/clinician survey as retrospective location of some data elements may be limited. However,</p>	<p>2g</p> <p>C <input type="checkbox"/></p> <p>P <input type="checkbox"/></p> <p>M <input type="checkbox"/></p> <p>N <input type="checkbox"/></p> <p>NA <input type="checkbox"/></p>

Comment [KP16]: 2e. For outcome measures and other measures (e.g., resource use) when indicated:
 •an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care) and are present at start of care; OR
 rationale/data support no risk adjustment.

Comment [k17]: 13 Risk models should not obscure disparities in care for populations by including factors that are associated with differences/inequalities in care such as race, socioeconomic status, gender (e.g., poorer treatment outcomes of African American men with prostate cancer, inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than adjusting out differences.

Comment [KP18]: 2f. Data analysis demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful differences in performance.

Comment [k19]: 14 With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74% v. 75%) is clinically meaningful; or whether a statistically significant difference of \$25 in cost for an episode of care (e.g., \$5,000 v. \$5,025) is practically meaningful. Measures with overall poor performance may not demonstrate much variability across providers.

Comment [KP20]: 2g. If multiple data sources/methods are allowed, there is demonstration they produce comparable results.

paper records should have some information and may be used to support the survey.	
2h. Disparities in Care	
2h.1 If measure is stratified, provide stratified results (<i>scores by stratified categories/cohorts</i>): N/A	2h C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/>
2h.2 If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans: N/A	
TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for <i>Scientific Acceptability of Measure Properties</i> ?	2
Steering Committee: Overall, to what extent was the criterion, <i>Scientific Acceptability of Measure Properties</i> , met? Rationale:	2 C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (evaluation criteria)	Eval Rating
3a. Meaningful, Understandable, and Useful Information	
3a.1 Current Use: testing not yet completed	
3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large) (<i>If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years</i>): Not currently reported publicly. However, appropriate use measures are being integrated into laboratory accreditation standards and as a result will be required for health insurance reimbursement.	
3a.3 If used in other programs/initiatives (<i>If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years</i>): Measures of appropriate use for cardiac imaging will be used as a part of an ACC National Quality Improvement Innovation Community called IMAGING in FOCUS (Formation of Optimal Cardiovascular IMAGING Utilization Strategies) to be launched in January 2010. www.acc.org/auc .	
Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)	
3a.4 Data/sample (<i>description of data/sample and size</i>): Limited use for QI in Hendel, RC; Cerqueira, M; Douglas, PS et al. "A Multicenter Assessment of the Use of Single-Photon Emission Computed Tomography Myocardial Perfusion Imaging With Appropriateness Criteria". J Am Coll Cardiol. Published online December 10, 2009.	
3a.5 Methods (<i>e.g., focus group, survey, QI project</i>): QI Project. Limited use for QI in Hendel, RC; Cerqueira, M; Douglas, PS et al. "A Multicenter Assessment of the Use of Single-Photon Emission Computed Tomography Myocardial Perfusion Imaging With Appropriateness Criteria". J Am Coll Cardiol. Published online December 10, 2009.	
3a.6 Results (<i>qualitative and/or quantitative results and conclusions</i>): The single site that had a substantial change in the rate of inappropriate test use initiated an internal analysis of appropriateness data and held group meetings and discussions to educate physicians on compliance with the AUC. The management team at this practice was highly motivated to improve performance and made education of physicians a priority; their overall inappropriate testing rate was the highest of all the sites at baseline and decreased from 22.0% to 13.3% at the end of study (p = 0.04).	3a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
3b/3c. Relation to other NQF-endorsed measures	

Comment [KP21]: 2h. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender); OR rationale/data justifies why stratification is not necessary or not feasible.

Comment [KP22]: 3a. Demonstration that information produced by the measure is meaningful, understandable, and useful to the intended audience(s) for both public reporting (e.g., focus group, cognitive testing) and informing quality improvement (e.g., quality improvement initiatives). An important outcome that may not have an identified improvement strategy still can be useful for informing quality improvement by identifying the need for and stimulating new approaches to improvement.

3b.1 NQF # and Title of similar or related measures:	
(for NQF staff use) Notes on similar/related <u>endorsed</u> or submitted measures:	
3b. Harmonization If this measure is related to measure(s) already <u>endorsed by NQF</u> (e.g., same topic, but different target population/setting/data source <u>or</u> different topic but same target population): 3b.2 Are the measure specifications harmonized? If not, why?	3b C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/>
3c. Distinctive or Additive Value 3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF- endorsed measures: 5.1 Competing Measures If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), describe why it is a more valid or efficient way to measure quality:	3c C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for <i>Usability</i> ?	3
Steering Committee: Overall, to what extent was the criterion, <i>Usability</i> , met? Rationale:	3 C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
4. FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (evaluation criteria)	Eval Rating
4a. Data Generated as a Byproduct of Care Processes 4a.1-2 How are the data elements that are needed to compute measure scores generated? <i>data generated as byproduct of care processes during delivery, Survey, coding/abstraction performed by someone other than person obtaining original information,</i>	4a C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
4b. Electronic Sources 4b.1 Are all the data elements available electronically? (<i>elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims</i>) No 4b.2 If not, specify the near-term path to achieve electronic capture by most providers. <i>Some data elements should already be a part of the electronic record (PCI history, scheduled surgery). In addition, e-ordering for diagnostic testing has been proposed for meaningful use, encouraging integration of these types of data elements. In addition, ACC is developing clinical decision support tools that can be embedded in electronic health records to capture the necessary information.</i>	4b C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
4c. Exclusions 4c.1 Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications? No 4c.2 If yes, provide justification.	4c C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/> NA <input type="checkbox"/>
4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences	4d

Comment [KP23]: 3b. The measure specifications are harmonized with other measures, and are applicable to multiple levels and settings.

Comment [k24]: 16 Measure harmonization refers to the standardization of specifications for similar measures on the same topic (e.g., *influenza immunization* of patients in hospitals or nursing homes), or related measures for the same target population (e.g., eye exam and HbA1c for *patients with diabetes*), or definitions applicable to many measures (e.g., age designation for children) so that they are uniform or compatible, unless differences are dictated by the evidence. The dimensions of harmonization can include numerator, denominator, exclusions, and data source and collection instructions. The extent of harmonization depends on the relationship of the measures, the evidence for the specific measure focus, and differences in data sources.

Comment [KP25]: 3c. Review of existing endorsed measures and measure sets demonstrates that the measure provides a distinctive or additive value to existing NQF-endorsed measures (e.g., provides a more complete picture of quality for a particular condition or aspect of healthcare).

Comment [k26]: 5. Demonstration that the measure is superior to competing measures - new submissions and/or endorsed measures (e.g., is a more valid or efficient way to measure).

Comment [KP27]: 4a. For clinical measures, required data elements are routinely generated concurrent with and as a byproduct of care processes during care delivery. (e.g., BP recorded in the electronic record, not abstracted from the record later by other personnel; patient self-assessment tools, e.g., depression scale; lab values, meds, etc.)

Comment [KP28]: 4b. The required data elements are available in electronic sources. If the required data are not in existing electronic sources, a credible, near-term path to electronic collection by most providers is specified and clinical data elements are specified for transition to the electronic health record.

Comment [KP29]: 4c. Exclusions should not require additional data sources beyond what is required for scoring the measure (e.g., numerator and denominator) unless justified as supporting measure validity.

Comment [KP30]: 4d. Susceptibility to inaccuracies, errors, or unintended consequences and the ability to audit the data items to detect such problems are identified.

<p>4d.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results. Submission of individual clinical data variables required for Framingham risk (ATP III criteria) calculation for asymptomatic patients is recognized to place a significant data collection burden upon institutions and may not be possible based on data elements that are readily available at the imaging laboratory. As such, a clinician estimate of CHD risk will be collected for all asymptomatic patients who are being seen for initial detection and risk assessment without known coronary heart disease. However, in making their estimate, clinicians should consider the maximum number of available patient factors used to estimate risk based on Framingham (ATP III criteria), typically age, gender, diabetes, smoking status, and use of blood pressure medication, and integrate age appropriate estimates for missing elements, such as LDL or standard blood pressure. While calculation of the estimate does not require submission of the actual clinical data elements other than the clinician estimate of CHD risk, clinicians are attesting to the accuracy of the estimate by submitting it. An audit of clinician estimates should be completed on a subset of clinicians to verify their estimates as being accurate based on the data that was available. The ACC/UHC SPECT MPI AUC pilot study did audit physician estimates and did find an overestimation of risk, primarily between low and intermediate risk. However, physicians in the study were not required to use all available data to make their estimate nor were they directly subject to individual audits.</p>	C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
<p>4e. Data Collection Strategy/Implementation</p> <p>4e.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: Hendel, RC; Cerqueira, M; Douglas, PS et al. "A Multicenter Assessment of the Use of Single-Photon Emission Computed Tomography Myocardial Perfusion Imaging With Appropriateness Criteria". J Am Coll Cardiol. Published online December 10, 2009.</p> <p>This study demonstrated the feasibility of data collection as well as the most frequent inappropriate indications. This allowed ACC to narrow the number of indications measured for this measure set along with the associated data elements. The time and data availability for Framingham risk calculation also changed our approach from actual calculation to an estimate.</p> <p>4e.2 Costs to implement the measure (costs of data collection, fees associated with proprietary measures): No data available, although the data elements required should be collected as a part of patient intake for testing.</p> <p>4e.3 Evidence for costs: None available.</p> <p>4e.4 Business case documentation: Given the expense of cardiovascular imaging, potential reductions in inappropriate test ordering should yield significant cost savings to the healthcare system.</p>	4e C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
<p>TAP/Workgroup: What are the strengths and weaknesses in relation to the sub-criteria for Feasibility?</p>	4
<p>Steering Committee: Overall, to what extent was the criterion, Feasibility, met? Rationale:</p>	4 C <input type="checkbox"/> P <input type="checkbox"/> M <input type="checkbox"/> N <input type="checkbox"/>
RECOMMENDATION	
<p>(for NQF staff use) Check if measure is untested and only eligible for time-limited endorsement.</p>	Time-limited <input type="checkbox"/>
<p>Steering Committee: Do you recommend for endorsement?</p>	Y <input type="checkbox"/>

Comment [KP31]: 4e. Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, etc.) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).

Comments:	<input type="checkbox"/> N <input type="checkbox"/> A
CONTACT INFORMATION	
Co.1 Measure Steward (Intellectual Property Owner) Co.1 Organization American College of Cardiology Foundation 2400 N St. NW Washington District Of Columbia 20037	
Co.2 Point of Contact Joseph Allen, MA jallen@acc.org 202-375-6463	
Measure Developer If different from Measure Steward Co.3 Organization American College of Cardiology Foundation 2400 N St. NW Washington District Of Columbia 20037	
Co.4 Point of Contact Joseph Allen, MA jallen@acc.org 202-375-6463	
Co.5 Submitter If different from Measure Steward POC Joseph Allen, MA jallen@acc.org 202-375-6463- American College of Cardiology Foundation	
Co.6 Additional organizations that sponsored/participated in measure development	
ADDITIONAL INFORMATION	
Workgroup/Expert Panel involved in measure development Ad.1 Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development. All individuals are volunteer members representing American College of Cardiology Foundation: Pamela Douglas, MD, MACC Joseph Allen, MA Robert Hendel, MD, FACC Joseph Cacchione, MD, FACC Manuel Cerqueira, MD, FACC Joseph Drozda, MD, FACC Michael Picard, MD, FACC Martha Radford, MD, FACC Leslee Shaw, PhD, FACC Allen Taylor, MD, FACC	
Ad.2 If adapted, provide name of original measure: Ad.3-5 If adapted, provide original specifications URL or attachment	
Measure Developer/Steward Updates and Ongoing Maintenance Ad.6 Year the measure was first released: 2009 Ad.7 Month and Year of most recent revision: 2009-12 Ad.8 What is your frequency for review/update of this measure? Annual Ad.9 When is the next scheduled review/update for this measure? 2010-12	
Ad.10 Copyright statement/disclaimers: Copyright 2009. American College of Cardiology Foundation	
Ad.11 -13 Additional Information web page URL or attachment:	
Date of Submission (MM/DD/YY): 05/14/2010	

**American College of Cardiology Foundation
Optimization of Patient Selection for Cardiac Stress Imaging
Data Collection Form**

Test Date: ___ / ___ / ___

Symptom Status

- Asymptomatic
- Ischemic Equivalent
- Not known/not available

Presence of Prior Known CHD

- Yes
- No
- Not known/not available

Risk Category OR Procedure Documentation

Prior Percutaneous Coronary Intervention (PCI)

Date of most recent PCI: ___ / ___ / ___ Date not available

Preoperative evaluation

Name of scheduled surgery: _____

Urgency of scheduled surgery

- Urgent
- Elective
- Not recorded/not available

Asymptomatic patient

Estimated CHD Risk

- Low CHD risk
- Intermediate CHD risk
- High CHD risk
- Not recorded/not available

Prior CHD assessment by one of the following methods no matter the test result:

- Exercise stress treadmill
- Non-invasive imaging
 - Stress echo
 - Stress SPECT MPI
 - CT Angiography
 - Calcium Scoring
- Invasive imaging (cardiac catheterization)
- No record of prior CHD assessment by one of the above methods

None of the above/not recorded

**American College of Cardiology Foundation
Optimization of Patient Selection for Cardiac Stress Imaging
Proposed Measurement Micro-Specifications
Version: Executive Committee Review**

All Measures

Data collection period:

Measures are to be developed based on a sample of a full calendar year based on the following sampling methodology:

- Select a starting month:
 - January
 - March
 - May
 - July
 - September
 - November
- Begin 60 day data collection period on the 1st on the month for the selected starting month
- Determine whether at least 30 stress SPECT and stress echo orders have been placed during the selected time period. If not, select another time period with a minimum number of 30 cases. If no time period includes the minimum number of cases, then the imaging laboratory does not have sufficient volume to report this measure.

Measure #1: Adequacy of data to assess appropriate use of cardiac stress imaging

The following definitions should be used in documenting the data to assess appropriate use of cardiac stress imaging:

Measure #1, Element #1: Symptom Status

Definition: Ischemic Equivalent: Chest Pain Syndrome, Anginal Equivalent, or Ischemic ECG Abnormalities: Any constellation of clinical findings that the physician feels is consistent with obstructive CAD. Examples of such findings include, but are not exclusive to, chest pain, chest tightness, burning, shoulder pain, palpitations, jaw pain, and new ECG abnormalities suggestive of ischemic heart disease. Non-chest pain symptoms, such as dyspnea or worsening effort tolerance, that are felt to be consistent with CAD may also be considered to be an anginal equivalent.

Data variables:

- Asymptomatic – not meeting definition of ischemic equivalent
- Ischemic Equivalent – meeting definition of ischemic equivalent

Measure #1, Element #2

- Known CAD, including
 - prior MI
 - prior ACS
 - prior CABG
 - prior PCI or
 - CHD on prior diagnostic test
 - Exercise stress treadmill
 - Non-invasive imaging
 - Stress echo
 - Stress SPECT MPI
 - CT Angiography
 - Calcium Scoring
 - Invasive imaging (cardiac catheterization)

Measure #1, Element #3a: Documentation of prior PCI at time of test requisition

If prior PCI, time since most recent PCI

Record the date (month, day, year) of the most recent PCI. If day, not known month and year is acceptable.

Measure #1, Element #3b: If perioperative evaluation, scheduled surgery

Patient is being seen for preoperative evaluation if:

- an upcoming surgery is the recorded reason for the imaging test AND
- no other reason is recorded for the imaging

The following information should be recorded:

1. name of the scheduled surgery
2. urgency of the scheduled surgery

The following categories will be used in Measure #4 to determine the risk of the surgery recorded for this element.

Surgical Risk Categories

- High-Risk Surgery– cardiac death or MI greater than 5%

Emergent major operations (particularly in the elderly), aortic and peripheral vascular surgery, prolonged surgical procedures associated with large fluid shifts and/or blood loss.

- Intermediate-Risk Surgery– cardiac death or MI equal to 1% to 5%

Carotid endarterectomy, head and neck surgery, surgery of the chest or abdomen, orthopedic surgery, prostate surgery.

- Low-Risk Surgery– cardiac death or MI less than 1%

Endoscopic procedures, superficial procedures, cataract surgery, breast surgery.

Measure #1, Element #3c: If initial risk assessment in asymptomatic patient,

A patient does **NOT** have:

- Known CAD, including
 - prior MI
 - prior ACS
 - prior CABG
 - prior PCI or
 - CHD on prior diagnostic test
 - Exercise stress treadmill
 - Non-invasive imaging
 - Stress echo
 - Stress SPECT MPI
 - CT Angiography
 - Calcium Scoring
 - Invasive imaging (cardiac catheterization)
- Ischemic equivalent
- Undergone prior CHD assessment by one the following methods no matter the test result:
 - Exercise stress treadmill
 - Non-invasive imaging
 - Stress echo
 - Stress SPECT MPI
 - CT Angiography
 - Calcium Scoring
 - Invasive imaging (cardiac catheterization)
- Preoperative evaluation

Asyptomatic patients should have recorded a clinician estimate of coronary heart disease risk category (ATP III criteria) based on the following methodology:

Clinical Estimate of Coronary Heart Disease Risk Category (ATP III criteria)

In making the clinical estimate of coronary heart disease (CHD) risk, clinicians should consider the maximum number of available patient factors used to estimate Framingham (ATP III criteria), typically age, gender, diabetes, smoking status, and use of blood pressure medication, and integrate age appropriate estimates for missing elements, such as LDL or standard blood pressure. While calculation of the estimate does not require submission of the actual clinical data elements other than the clinician estimate of CHD risk, clinicians are attesting to the accuracy of the estimate by submitting it. An audit of clinician estimates should be completed on a subset of clinicians to verify their estimates as being accurate based on the data that was available.

Absolute CHD risk is defined as the probability of developing CHD, including myocardial infarction or CHD death over a given time period. The ATP III report

specifies absolute risk for CHD over the next 10 years. CHD risk refers to 10-year risk for any hard cardiac event (mortality and myocardial infarction).

- **CHD Risk—Low**

Defined by the age-specific risk level that is below average. In general, low risk will correlate with a 10-year absolute CHD risk less than 10%.

- **CHD Risk—Moderate**

Defined by the age-specific risk level that is average or above average. In general, moderate risk will correlate with a 10-year absolute CHD risk between 10% and 20%.

- **CHD Risk—High**

Defined as the presence of diabetes mellitus in a patient 40 years of age or older, peripheral arterial disease or other coronary risk equivalents, or a 10-year absolute CHD risk of greater than 20%.

National Heart, Lung, and Blood Institute report on "[Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults \(Adult Treatment Panel III\)](#)" (ATP III).

Measure #2: Cardiac stress imaging not meeting appropriate use criteria: Testing in asymptomatic, low risk patients

- For all orders in asymptomatic patients, determine orders for initial diagnosis and risk assessment. In doing so, patients with known CHD, prior PCI or prior CABG and the following exclusions are not included.

Patients qualify for this measure if:

- Asymptomatic AND
- Low CHD risk based on physician estimate AND

NOT any of the following:

- Known CAD, including
 - prior MI
 - prior ACS
 - prior CABG
 - prior PCI or
 - CHD on prior diagnostic test
 - Exercise stress treadmill
 - Non-invasive imaging
 - Stress echo
 - Stress SPECT MPI
 - CT Angiography
 - Calcium Scoring
 - Invasive imaging (cardiac catheterization)
- Ischemic equivalent
- Undergone prior CHD assessment by one the following methods no matter the test result:
 - Exercise stress treadmill
 - Non-invasive imaging
 - Stress echo
 - Stress SPECT MPI
 - CT Angiography
 - Calcium Scoring
 - Invasive imaging (cardiac catheterization)
- Patients for whom preoperative testing is the primary reason for imaging

Measure #3: Cardiac stress imaging not meeting appropriate use criteria: Routine testing after percutaneous coronary intervention (PCI)

- For all orders post PCI, determine all orders that were in asymptomatic patients
- Among asymptomatic patients, subtract date of most recent PCI from date of test requisition and categorize into orders less than two years since most recent PCI and orders placed greater than or equal to two years since most recent PCI

Patients qualify for this measure if:

- Asymptomatic AND
- Less than two years since most recent PCI

Measure #4: Cardiac stress imaging not meeting appropriate use criteria: Preoperative evaluation in low risk surgery patients

Patient qualify this measure if:

- an upcoming surgery is the recorded reason for the imaging test AND
- no other reason is recorded for the imaging

AND

Surgery risk is low

The following categories will be used to determine whether the risk of the surgery recorded is low:

Surgical Risk Categories

- Low-Risk Surgery– cardiac death or MI less than 1%
Endoscopic procedures, superficial procedures, cataract surgery, breast surgery.

Examples of Low Risk Surgeries

Cataract laser surgery
Bx breast percut w/device
Clos Large Bowel Biopsy
Closed Bronchial Biopsy
Colonoscopy
Colonoscopy
Colonoscopy
Colonoscopy and biopsy
Colonoscopy and biopsy
Colonoscopy, lesion removal
Colonoscopy, lesion removal
Cystoscopy
Cystoscopy
Cystoscopy Nec
Diagnostic colonoscopy
Diagnostic colonoscopy
Diagnostic laryngoscopy
Egd With Closed Biopsy
Endo Polpectomy Lrge Int
Esophagus endoscopy,dilation
Intraoper Cholangiogram
Nasal endoscopy, dx
Percu Endosc Gastrostomy (Begin 1989)
Sm Bowel Endoscopy Nec
Upper GI endoscopy, biopsy
Upper GI endoscopy, biopsy

Upper gi endoscopy,diagnosis
Upper gi endoscopy,diagnosis

Surgeries classified as low risk should NOT meet the following definitions.

- Intermediate-Risk Surgery– cardiac death or MI equal to 1% to 5%
Carotid endarterectomy, head and neck surgery, surgery of the chest or abdomen, orthopedic surgery, prostate surgery.
- High-Risk Surgery– cardiac death or MI greater than 5%
Emergent major operations (particularly in the elderly), aortic and peripheral vascular surgery, prolonged surgical procedures associated with large fluid shifts and/or blood loss.

<u>Measure #/Title/Steward</u>	
#IEP-016-10/ Cardiac stress imaging not meeting appropriate use criteria: Testing in asymptomatic, low risk patients/ACC	
<u>Description</u>	
Percentage of all stress SPECT MPI and stress echo performed in asymptomatic, low CHD risk patients for initial detection and risk assessment.	
<u>Initial In-person Vote</u>	
THE STEERING COMMITTEE DID NOT VOTE ON THE MEASURE AT THE TIME OF THE IN-PERSON MEETING – THE STEERING COMMITTEE WILL VOTE ON THE MEASURE AFTER THE MEASURE DEVELOPER HAS RESPONDED TO THE CONDITIONS FOR RECOMMENDATION.	
Steering Committee Questions/Conditions for Measure Developer:	
Abbreviated Response from Measure Developer:	
<ul style="list-style-type: none"> Measure as it stands needs to consider the need for risk adjustment or provide rationale for no risk adjustment 	<ul style="list-style-type: none"> The measure is specified to account for risk. It accounts for the patient’s clinical characteristics and includes mandatory exclusions from the measured population.
<ul style="list-style-type: none"> Expand to include MRI or provide a justification for not including MRI in the scope of the measure 	<ul style="list-style-type: none"> ACC is willing to add CTA and MRI to the measure. <ul style="list-style-type: none"> Note: ACC does not anticipate a large number of cases to be documented for these imaging modalities in this patient population.
<ul style="list-style-type: none"> Remove the denominator exclusion criteria, “patients without sufficient patient selection criteria recorded” 	<ul style="list-style-type: none"> ACC is willing to meet condition, but noted the change may inflate the denominator and have potential unintended consequences
<ul style="list-style-type: none"> Consider changing the title of the measure, removing potentially negative connotations. 	<ul style="list-style-type: none"> Does not support. ACC elected not to meet this condition, rational – not warranted
Detailed Response from Measure Developer:	
<ul style="list-style-type: none"> Risk is explicitly considered in the measures as it takes into account two clinical characteristics of the patient – symptom status and global risk for CHD. The latter includes numerous factors including age, gender, smoking status, blood pressure, lipid profile etc. Exclusions for a known history of CHD, periop evaluation, and prior testing also are included to ensure that patients who are not being seen for initial evaluation of CHD are excluded. Additional risk adjustment is not required since patient risk is already core to the definition of this measure. ACC is willing to add CTA and MRI to the measure. However, ACC does not anticipate a large number of cases to be documented for these imaging modalities in this patient population. ACC is willing to remove the denominator exclusion criteria, “patients without sufficient patient selection criteria recorded.” However, this change will inflate the denominator of the measure for imaging laboratories that are unable to locate the information necessary to determine all components of the numerator. As such, the removal could create an incentive not to obtain enough data to clearly indicate if 	

a patient qualifies for this measure.

- ACC does not think a change of the title to be less negative is warranted. The College revised its initial draft measure title of “Inappropriate cardiac stress imaging” to the current title in an attempt to make the title more neutral. The measure is designed to examine imaging that is not reasonable. Development of a more positive title would not reflect the focus of this measure which is overuse.
-