MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 0066
De.2. **Measure Title:** Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%)
Co.1.1. **Measure Steward:** American Heart Association
De.3. **Brief Description of Measure:** Percentage of patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12 month period who also have diabetes OR a current or prior Left Ventricular Ejection Fraction (LVEF) < 40% who were prescribed ACE inhibitor or ARB therapy
1b.1. **Developer Rationale:** In the absence of contraindications, ACE inhibitors or ARBs are recommended for all patients with a diagnosis of coronary artery disease and diabetes or reduced left ventricular systolic function. ACE inhibitors remain the first choice, but ARBs can now be considered a reasonable alternative. Both pharmacologic agents have been shown to decrease the risk of death, myocardial infarction, and stroke. Additional benefits of ACE inhibitors include slowed disease progression and reduction of complications for patients with diabetes.

S.4. **Numerator Statement:** Patients who were prescribed ACE inhibitor or ARB therapy
S.7. **Denominator Statement:** All patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12 month period who also have diabetes OR current or prior LVEF <40%
S.10. **Denominator Exclusions:** Documentation of medical reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, allergy, intolerance, pregnancy, renal failure due to ACE inhibitor, diseases of the aortic or mitral valve, other medical reasons)

Documentation of patient reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, patient declined, other patient reasons)

Documentation of system reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, lack of drug availability, other reasons attributable to the health care system)

De.1. **Measure Type:** Process
S.23. **Data Source:** Electronic Clinical Data : Registry
S.26. **Level of Analysis:** Clinician : Group/Practice, Clinician : Individual

IF Endorsement Maintenance – Original Endorsement Date: Aug 10, 2009 **Most Recent Endorsement Date:** Jan 18, 2012

Maintenance of Endorsement -- Preliminary Analysis

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

1a. Evidence. The evidence requirements for a *process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Summary of prior review in 2012

- The developer included the steps between the diagnosis of patients with CAD and diabetes or LVSD with a prescription for an ACE or ARB and the reduced risk of death, MI, and stroke.
- The developer provided one clinical guideline with two recommendations from the 2007 chronic angina focused update of the ACC/AHA 2002 Guidelines for the Management of Patients with Chronic Stable Angina:
  - ACE inhibitors should be started and continued indefinitely in all patients with left ventricular ejection fraction less than or equal to 40% and in those with hypertension, diabetes, or chronic kidney disease, unless contraindicated. **Level of Evidence: Level A**
  - Angiotensin receptor blockers are recommended for patients who have hypertension, have indicators for but are intolerant of ACE inhibitors, have heart failure, or have had a myocardial infarction with left ventricular ejection fraction less than or equal to 40%. **Level of Evidence: Level A**

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- The developer provided an additional guideline with two recommendations for renin-angiotensin-aldosterone blocker therapy from the 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the Diagnosis and Management of Patients with Stable Ischemic Heart Disease:
  - ACE inhibitors should be prescribed in all patients with stable ischemic heart disease (SIHD) who also have hypertension, diabetes mellitus, LVEF 40% or less, or chronic kidney disease (CKD), unless contraindicated. **Level of Evidence: Level A**
  - ARBs are recommended for patients with stable ischemic heart disease (SIHD) who have hypertension, diabetes mellitus, LV systolic dysfunction, or chronic kidney disease (CKD) and have indications for, but are intolerant of, ACE inhibitors." **Level of Evidence: Level A**
- The developer provided a systematic review of the body of evidence supporting the benefits of ACE inhibitor/ARB therapy for patients with ischemic heart disease and included a summary of the Quantity, Quality, and Consistency of the body of evidence.

Exception to evidence:

N/A

Guidance from the Evidence Algorithm: Process measure with SR and grading of the body of evidence (Box 3) →Summary of QQC (Box 4) →SR concludes QQC is High (Box 5a) → High

2012 discussion: *In 2012, during the previous review of this measure, the Steering Committee questioned why patients with CAD and HTN, or patients with CAD and CKD were not included in the measure. At the time, the developer responded, the guidelines did not explicitly recommend ARBs for patients with HTN or CKD.*

Questions for the Committee:

- *Is the Committee willing to accept the prior evaluation? The updated evidence supports the measure focus and includes patients that were not previously included.*

Preliminary rating for evidence: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities
Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- The developer provided the following average performance rates from the PQRS Experience Report from 2011 – 2014:

Year	Average Performance Rate
2011	63.5%
2012	64.0%
2013	70.0%
2014	81.2%

- For endorsement maintenance, NQF asks for performance scores (current and over time), including mean, standard deviation, min, max, interquartile range, scores by decile, and a description of the data source (number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included).
- The developer also provided the following summary of NCDR PINNACLE Registry data from the literature:

Year(s)	Description	Prescription Rate	Median Rx Rate	Range
2009 - 2012	Rx rates of ACEI/ARBs for pts w/CAD & concurrent DM or LVSD	69.6% - 77.6%	--	--
2008 - 2010	Outpatient cardiology facilities prescribing ACEI/ARBs for pts. w/ CAD at index clinic visit	69.4% (index visit)	75.5%	39.1%-100%
		72.3% (w/in yr. of index visit)	78.1%	45.4%-100%
2008-2009	ACEI/ARB Rx among CAD pts w/concurrent LVSD or DM	72.4%	--	--

Disparities:

- The developer did not provide any data on disparities from the measure as specified – this is encouraged for endorsement maintenance.
- The developer stated that while this measure is included in federal reporting programs, those programs have not yet made disparities data available to analyze and report.
- The developer provided data on disparities from Chan et al. (2010) using PINNACLE Registry data from 2009 that demonstrated a slight difference between men and women (72.1% vs. 71.7%) with LVSD or diabetes who were prescribed ACEI/ARBs.
- A separate analysis of 2009 PINNACLE Registry data by Smolderen (2013) evaluated the impact of CAD patients' insurance status and the likelihood they would be prescribed an ACEI/ARB:

Insurance Coverage	Prescription rate
Privately-insured	75.5%
Publicly-insured	69.1%
Uninsured	66.7%

Questions for the Committee:

- *Does the data presented adequately demonstrate that there is a quality problem in prescribing ACEI/ARB therapy for patients with CAD and diabetes or heart failure?*
- *Is there an opportunity for improvement in quality that warrants the continued endorsement of a national performance measure?*
- *Are you aware of evidence that additional disparities exist in this area of healthcare?*

Preliminary rating for opportunity for improvement: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1. Importance to Measure and Report

1a. Evidence to Support Measure Focus

Comments: **"The evidence directly related to the process. ACEs and ARBs improve outcomes in the denominator population.

I am not aware of any studies that refute the evidence"

**High Evidence

** Several strong randomized controlled trials and two meta-analyses supporting it. The evidence appears to be a class effect of ACE/ARB and there is emerging data that angiotensin receptor-neprilysin inhibitor (ARNI) have similar effects in the HF population (SICD not yet defined that I am aware of). The evidence appears to directly support ACE/ARB use in these populations. Canadian guidelines have already shifted towards use of ARNI as well in HF populations. The process is supported by the rationale.

**This process measure is based on systematic review of the evidence and grading where specific focus matches what is being measured. They included the steps between diagnosis of patients with CAD and diabetes or LVSD with prescription for ACE or ARB and reduced risk of death, MI and stroke.

**"The PARADIGM study is not included in this measure review. That study found a 22% RRR in mortality and hospitalization of patients receiving an ARNI (Angiotensin receptor and Nephrilyn Inhibitor) over those receiving enalapril. This information changed the ACCF/AHA guidelines update published in May, 2016 to include this drug as a Class I recommendation for those with a reduced EF. So, it should be considered for the person with an MI and reduced Ef, but not for the post MI with DM.

Also this year a drug was approved by the FDA that can be used to treat hyperkalemia (Valtessa). It can be used to treat chronic hyperkalemia and thus allow RAAS inhibitors to be continued to uptitrated. The OPAL, PEARL, and AMYTHIST trials have demonstrated 1 month and 12 month efficacy. The drug has difficulty in clinical use due to dosing problems. It is worthy of discussion as it might have an influence on exclusions for this measure.

This process measure was initially approved in 2009 and was re-approved in 2012. It has long standing supportive evidence for inclusion. Average performance rate has continued to improve and was at 81% in 2014. but there remains enough room for growth that this should be retained.

**The studies relate directly to the focus of the measure

1b. Performance Gap

Comments: **"2011: 63.5%

2012: 64.0%

2013: 70.0%

2014: 81.2%

A significant gap continues to exist. The developer calculated measures of tendency, variability, and dispersion based on the sample of 1,128 physicians:

- Mean performance rate: 0.71
- Median performance rate: 0.74
- Mode: 0.67
- Standard deviation: 0.19
- Min, Max: 0.00-1.00
- IQR: 0.18 (0.64 – 0.83)

**Moderate

**PQRS experience report has been provided from 2014 with improving performance over time (voluntary physician reporting

program). One would expect higher scores in present day if quality remained steady or improved to present day. In addition, there is a broad opportunity for improvement in this area also taking into consideration that practices who are not currently reporting likely have considerably lower ratings. Disparity data is available and is preliminary but shows there are no racial disparities but there may be some disparity in economic status due to patients being insured or uninsured.
**From the PQRS Experience Report from 2011-2014 shows a year to year increase in average performance rate with an increase of almost 18% in 4 years. That's good but still needs improvement. They look at disparities with small gap between men and women, larger between private/publicly insured but don't touch socioeconomic factors or language barriers.
**"As stated above, there is support for retaining (and perhaps refining) the measure) as there are disparities based on insurance.
**While there is still a gap in care, the gap has closed 20% in the our years between 2011 and 2014. So that the importance is much less than it was in the past.
1c. High Priority (previously referred to as High Impact)
Comments: **N/A
** N/A
**The composite performance measure has high quality construct and is logical. It should positively affect a large population due to a disease with high incidence/prevalence (and likely increasing) in reducing both morbidity from SICD and mortality. The measure appears to be high priority taking this into consideration as well as economic impact upon the US healthcare system.
**There is certainly room for improvement in overall quality of care. More emphasis on Gap in Care/Opportunity for Improvement and definitely on disparities in care.
**As stated the process measure is logical. We do need to address the most recent HF guidelines in relation to possible revisions.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

2a1. Reliability Specifications
Maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented.
Data source(s): electronic clinical registry data. This is not an eMeasure.
Specifications:

- The level of analysis is at the clinician-level.
- The numerator includes patients who were prescribed ACE inhibitor or ARB therapy.
  - 'Prescribed' may include an ACE or ARB prescription given to the patient at one or more visits in the 12 month measurement period OR a patient already taking an ACI or ARB as documented in their current medication list.
- The denominator includes all patients aged 18 and older with a diagnosis of CAD who are seen within a 12-month period who also have diabetes or current or prior LVEF <40%.
  - Two populations are included in the denominator and combined for a single reported performance score. If a patient has both DM and LVSD, reporting criteria #2 (CAD w/DM) will count as appropriate reporting for this patient.
    - Population 1- patients aged 18 and older with a diagnosis of CAD with LVEF <40%. LVEF < 40% corresponds to qualitative documentation of moderate or severe left systolic dysfunction.
    - Population 2- patients aged 18 and older with a diagnosis of CAD who have diabetes.
  - Two or more encounters are required to establish that the eligible professional has an existing relationship with the patient.
- The denominator exclusions included are:
  - Documentation of medical reason for not prescribing ACE inhibitor or ARB therapy (e.g., allergy, intolerance, pregnancy, renal failure due to ACE inhibitor, diseases of the aortic or mitral valve, other medical reasons)

- Documentation of patient reason(s) for not prescribing ACE inhibitor or ARB therapy (e.g., patient declined, other patient reasons)
- Documentation of system reason(s) for not prescribing ACE inhibitor or ARB therapy (e.g., lack of drug availability, other reasons attributable to the health care system)

- The ICD-9, ICD-10, CPT, and Report Quality Data codes have been included in the specification details.
- The calculation algorithm is included.
- The developer specifies how to handle missing data.

Questions for the Committee :

- *Are all the data elements clearly defined? Are all appropriate codes included?*
- *Is the logic or calculation algorithm clear?*
- *Is this measure consistently implemented?*

2a2. Reliability Testing attachment
Maintenance measures – less emphasis if no new testing data provided

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers.

For maintenance measures, summarize the reliability testing from the prior review:

- Inter-rater reliability was conducted on 100 randomly selected paper medical charts from four physician practices submitted to PQRI in 2007. The developer provided kappa statistics for seven data elements, however, only one data element was relevant to this measure – Diagnosis of CAD. The kappa statistic for 'Diagnosis of CAD' was 1.00.

Describe any updates to testing: The developer conducted measure score reliability testing – see below

SUMMARY OF TESTING

Reliability testing level ☒ **Measure score** ☐ **Data element** ☐ **Both**

Reliability testing performed with the data source and level of analysis indicated for this measure ☐ **Yes** ☒ **No**

Method(s) of reliability testing:

- The dataset included 2014 EHR data from PQRS provided by CMS. A total of 2,296 physicians reported on this measure in 2014. Of those, 1,128 (49.1%) physicians had all of the required data elements and a **minimum of 10 quality reporting events.** The **average number of quality reporting events was 49.0.** A total of 55,272 patients were included in the sample.
- The developers used a beta-binomial model to assess the signal-to-noise ratio. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in performance. The higher the reliability score, the greater is the confidence with which one can distinguish the performance of one physician from another. This is an appropriate test for measure score reliability. A reliability of 0.70 is generally considered a minimum threshold for reliability.

Results of reliability testing:

- Reliability at the minimum level of quality reporting events (10) was **0.58** and **0.87** at the average number of quality events (49.0).

Guidance from the Reliability Algorithm: Precise specifications (Box 1) → Empirical reliability testing (Box 2) → Computed performance scores for measured entities (Box 4) → Appropriate method used (Box 5) → Moderate

Questions for the Committee:

- *Is the measure score test sample adequate to generalize for widespread implementation?*
- *Do the results from the updated testing demonstrate sufficient reliability so that differences in performance between physicians can be identified?*

Preliminary rating for reliability: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

2b. Validity
Maintenance measures – less emphasis if no new testing data provided

2b1. Validity: Specifications

2b1. Validity Specifications. This section should determine if the measure specifications are consistent with the evidence.

Specifications consistent with evidence in 1a. ☒ **Yes** ☐ **Somewhat** ☐ **No**

Question for the Committee:

- *Are the specifications consistent with the evidence?*

2b2. Validity testing

2b2. Validity Testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

For maintenance measures, summarize the validity testing from the prior review:

- The developer stated that all PCPI measures were assessed for content validity by expert work group members during the development process. Additional input on content validity was obtained through public comment and a panel of consumer, purchaser, and patient representatives convened by PCPI.

Describe any updates to validity testing: The developer conducted face validity – see below

SUMMARY OF TESTING

Validity testing level ☒ **Measure score** ☐ **Data element testing against a gold standard** ☐ **Both**

Method of validity testing of the measure score:

- ☒ **Face validity only**
- ☐ **Empirical validity testing of the measure score**

Validity testing method:

- An panel comprised of 18 experts from the AHA Council on Clinical Cardiology systematically assessed whether the performance scores from the measure as specified could be used to distinguish good from poor quality. This is an appropriate method for face validity.

Validity testing results:

- **94.4%** (17) of the respondents either agreed or strongly agreed that this measure can accurately distinguish good and poor quality.

Questions for the Committee:

- *Do the results demonstrate sufficient validity so that conclusions about quality can be made?*
- *Do you agree that the score from this measure as specified is an indicator of quality?*

2b3-2b7. Threats to Validity

2b3. Exclusions:

- The developer listed two exceptions below, however, the specifications include three exceptions.
  - Documentation of medical reason(s) for not prescribing ACI/ARB therapy.
  - Documentation of patient reason(s) for not prescribing ACI/ARB therapy.
- Out of the 1,128 physicians (with minimum of 10 quality reporting events) there were a total of 2,222 exceptions.
  - Average number of exceptions per physician was 2.0.
  - Overall exception rate was 3.9%.
- It is not clear if the exceptions analysis includes all of the exceptions in the specifications or just the

two listed in this section.

Questions for the Committee:

- *Are any patients or patient groups inappropriately excluded from the measure?*
- *Are the exceptions of sufficient frequency and variation across providers to be needed (and outweigh the data collection burden)?*

2b4. Risk adjustment: **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

2b5. Meaningful difference *(can statistically significant and clinically/practically meaningful differences in performance measure scores can be identified)*:

- The developer calculated measures of tendency, variability, and dispersion based on the sample of 1,128 physicians:
  - Mean performance rate: 0.71
  - Median performance rate: 0.74
  - Mode: 0.67
  - Standard deviation: 0.19
  - Min, Max: 0.00-1.00
  - IQR: 0.18 (0.64 – 0.83)

Question for the Committee:

- *Does this measure identify statistically and clinically meaningful differences about quality among providers?*

2b6. Comparability of data sources/methods:

- Measure is not specified for more than one data source; comparability of data sources is not needed.

2b7. Missing Data

- The developer stated that data are not available to complete an analysis of missing data but did specify how missing data are handled.

Guidance from Validity Algorithm: Specifications consistent with evidence (Box 1) → Potential threats to validity assessed (Box2) → Empirical validity testing (Box 3)→Face validity of measure score systematically assessed (Box 4) → Agreement by expert panel that performance measure score distinguishes quality (Box 5)→Moderate (highest eligible rating is MODERATE)

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2d)

2a1. & 2b1. Specifications

Comments: **Specifications are clear.

**moderate

**Data elements, numerator and denominator appear clearly defined. Would suggest an external validation by NQF or other party of ICD9 and 10 codes. I have no concerns that this measure can be consistently implemented as long as the coding is correct. My only concern is if all similar drugs that show benefit in the SICD population, and if other populations can be included (e.g. CKD) in this measure in the future to broaden its benefit. No risk stratification is needed. Calculation of score is clearly defined as well as included and excluded population.

**The Inter-rater reliability testing was conducted on 100 randomly selected paper medical charts from four physician practices. The developer provided kappa statistics for seven data elements; only one was relevant to measure-diagnosis of CAD. The measure score is inadequate for widespread application and/or differences in physician performance.

**The data elements are clearly defined. The logic algorithm is clear and the measure can be consistely implemented.

**The reliability seems to be well defined.

**They only report on 2 of the 3 exclusions (medical and patient but not system)
**moderate
**Empirical validity testing did not occur. It may be beneficial in the future if Canadian or European data can be provided to evaluate the effects of similar measures in this population in the future. Face validity of the measure was assessed by an expert panel which demonstrated approximately 95% of the experts either agreed or strongly agreed that the measure can accurately distinguish good and poor quality. Although results may be confounded, it would externally validate the results if the Pinnacle registry or other data could be evaluated to determine if patients with and without ACE/ARB have the intended benefit and outcomes are indeed reduced in real world populations. This data may however be confounded, and would need to be adjusted for the possibility that patients not receiving ACE/ARB may not have appropriate documentation in their charts (or retrievable in discrete data) for the reasons why they are not receiving the ACE/ARB (e.g. cough, renal issues etc). Similar external validation may possibly be assessed through collaboration with non-US entities.
**I find this question confusing. Under validity, there are no specifications listed unless this refers to testing. In the evidence plus the introduction the purpose is to decrease the risk of death, MI, stroke, disease progression and complications. Whether the target population is the physician or the patient, the values and meaningful outcomes should be consistent.
**The validity comes from a consensus panel of 18, which may not be an adequate test of validity but there is face validity which has not been measured.

2a2. Reliability Testing

Comments: **Inter-rater reliability was conducted on 100 randomly selected paper medical charts from four physician practices submitted to PQRI in 2007. The developer provided kappa statistics for seven data elements, however, only one data element was relevant to this measure – Diagnosis of CAD. The kappa statistic for 'Diagnosis of CAD' was 1.00.
**moderate
**Reliability appears to improve with increase from minimum to average to high level of quality reporting events. Updated data would be beneficial if this can be provided.
**The dataset included a total of 2,296 physicians but only .49% included all the data elements. That is a very small sample size to base a national performance score on with consistently reliable results. The developers used a beta-binomial model to assess the signal-to-noise ratio. Reliability Test Scores vary from 0.58 to 0.87 with the minimum threshold considered reliable of 0.70.
**Reliability testing was moderate to low (58%) when only 10 quality reporting events was used; however, with the average reporting events of 49 it improved to .87. This is acceptable for widespread implementation. Agree with moderate reliability rating.
**the reliability is appropriate with the testing with 1128 physicians

2b2. Validity Testing

Comments: **The developer stated that all PCPI measures were assessed for content validity by expert work group members during the development process. Additional input on content validity was obtained through public comment and a panel of consumer, purchaser, and patient representatives convened by PCPI. 94.4% (17) of the respondents either agreed or strongly agreed that this measure can accurately distinguish good and poor quality.
**moderate
**Face validity was moderate with no empirical validity testing. If outcomes are captured in the Pinnacle or other AHA/ACC databases, it would be beneficial to evaluate outcomes stratified by patients who did and did not receive ACE/ARB (appropriately adjustments) either retrospectively or prospectively in the future to validate at a high level (Level 1) the proposed measure has the intended effect on this population. Assessment for any adverse outcomes would also be beneficial if possible.
**The developers stated all PCPI Measures were assessed by expert work group members with additional input through public comment, panel of consumer, purchaser and patient reps. Face validity by experts from AHA Council 94% agreed could distinguish good/poor quality.
**Face validity with a panel of experts was conducted with strong results. No empirical data to support this measure are available even though this measure has been in use for 6 years. Despite this relatively weak method for assuring validity, I support continued use of this measure but would encourage validation studies. There was adequate variability found in practice.

2b3. Exclusions Analysis

2b4. Risk Adjustment/Stratification for Outcome or Resource Use Measures

2b5. Identification of Statistically Significant & Meaningful Differences In Performance

2b6. Comparability of Performance Scores When More Than One Set of Specifications

2b7. Missing Data Analysis and Minimizing Bias

Comments: ** Exception rate wsa 3.9%. Whether systems exceptions were included was unclear. There was no risk adjustment Missing data - Patient eligibility is determined by a set of defined criteria relevant to a particular measure. If data required to determine patient eligibility are missing, those patients/cases would be ineligible for inclusion in the denominator and therefore the patient/case would be deleted. If data required to determine if a denominator eligible patient qualifies for the numerator (or has a valid exclusion/exception) are missing, this case would represent a quality failure.
**yes
**It does not appear there are any threats to validity other than those listed above. Overall exception rate appears quite low at less than 4%. It appears that the performance measure may actually be expanded to other populations in the future. No risk adjustment is needed or defined. High performance in the measure would indicate high quality with low rates of outcomes. Measures of central tendency, variability and dispersion demonstrate there was meaningful variation between practices/physicians. "
**From 1128 physicians, there were 2222 exceptions. The measure does not include disparities data or risk adjustment. We don't know if excluded persons have been unjustly excluded. We don't know if exceptions are of sufficient frequency and variation to outweigh the data collection burden. Was documentation of medical reasons for not prescribing ACI/ARB therapy. Was documentation of patient reasons for not taking ACI/ARB adequate? Was statistically significant differences meaningful? Are the data sources/methods comparable? Missing data was stated to be accounted for but method or sources not identified.
*2b3-7 Threats to Validity: It would be interesting to see what the medical reasons are if broad categories can be used. Perhaps worsening renal function or hypotension and other. For patient reasons it would be of interest to see if potential pregnancy is the reason for women to refuse the medication.
2b4 There is no risk adjustment
2b5. There is variability in scores, but no description of low vs high score practices.
2b6 NA
2b7 Information on missing data not available."

Criterion 3. Feasibility
Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- The data elements are generated during the routine delivery of care and available in defined fields in electronic clinical data.
- Commercial use of the measure requires a license agreement between the user and the AMA, ACC, or AHA. For noncommercial purposes, the measure (while copyrighted) can be reproduced and distributed, without modification.
- The developer did not provide information on the fees associated with the PINNACLE Registry or whether participation in the registry is limited to cardiologists.

Questions for the Committee:
- *Are the required data elements routinely generated and used during care delivery?*
- *Are the required data elements available in electronic form, e.g., EHR or other electronic sources?*

Preliminary rating for feasibility: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments
Criteria 3: Feasibility

3a. Byproduct of Care Processes
3b. Electronic Sources
3c. Data Collection Strategy
Comments: ** No concerns

**high
**High feasibility with discrete data and codes demonstrating a high capture of all necessary fields.
**Implementation of the measure may have issues. Are the data elements readily available in clearly defined fields? What is the licensing agreement between user and developer? What are the fees associated with the PINNACLE Registry and is participation limited to cardiologists.
**Data elements are routinely used and generated in the EHR.
**This is a very feasible measure using claims data

Criterion 4: Usability and Use
Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact /improvement and unintended consequences

4. Usability and Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

Current uses of the measure

Publicly reported? ☐ **Yes** ☒ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No**

Accountability program details:

- Physician Quality Reporting System (PQRS) (CMS): Beginning in 2015, the program applied a payment adjustment to EPs who do not satisfactorily report data on quality measures for covered professional services in 2013. The developer also stated that CMS announced that there are plans to make all PQRS individual EP level PQRS measures available for public reporting on Physician Compare in late 2017.
    - This measure has been endorsed since 2009 - per NQF criteria, performance results are used in at least 1 accountability application within 3 years after initial endorsement and are publicly reported within 6 years after initial endorsement (or the data on performance results are available).
- The PINNACLE Registry® is cardiology´s largest outpatient quality improvement registry, capturing data on coronary artery disease, hypertension, heart failure and atrial fibrillation. As of the fourth quarter of 2014, the registry has more than 28 million patient encounter records.

Improvement results:

- The developer included the performance rates previously reported in 1b.2. Progress on improvement, including trends in performance results, number and percentage of people receiving high-quality healthcare, geographic area and number and percentage of accountable entities and patients were not discussed.

Unexpected findings (positive or negative) during implementation: none listed

Potential harms:

- The developer stated that they are not aware of any unintended consequences .

Feedback :

- In 2012, the Steering Committee stated that this was an important clinical measure; however, a more stringent numerator criteria (i.e. must have X number of refills within defined time frame) would make it a stronger measure.

***Questions for the Committee**:*

- *How can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

Preliminary rating for usability and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 4: Usability and Use

4a. Accountability and Transparency

4b. Improvement

4c. Unintended Consequences

Comments: **Used by PQRS and NCDR Pinnacle Registry

**moderate

**High usability and use and a very important measure overall in improving US cardiovascular care. Endorsing the measure would likely result in reduced US healthcare costs and improved outcomes reducing both morbidity and mortality from important outcomes associated with SICD in the defined populations.

**The uses of the measure is not currently publicly reported but it is currently used in an accountability program. How will EPs be evaluated and fee adjusted? Will EP Accountability programs contribute to improved performance results, the number and percentage of people receiving high-quality healthcare and consistency across geographic/socioeconomic/ethnic groups?

**Not publicly reported at this time. It is used for performance improvement.

**Currently the measure is not being publicly reported.

Criterion 5: Related and Competing Measures

Related or competing measures:

- 0067 : Chronic Stable Coronary Artery Disease: Antiplatelet Therapy
- 0070 : Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)
- 0074 : Chronic Stable Coronary Artery Disease: Lipid Control
- 0081 : Heart Failure (HF): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy for Left Ventricular Systolic Dysfunction (LVSD)
- 1522 : ACE/ARB Therapy at Discharge for ICD implant patients with Left Ventricular Systolic Dysfunction
- 1662 : Angiotensin Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy
- 2467 : Adherence to ACEIs/ARBs for Individuals with Diabetes Mellitus

Harmonization:

- The developer stated that this measure's specifications are harmonized with existing measures where possible but there are several key differences:
  - NQF #1662, 1522, 0081, 2467: focus on the prescription of ACEI/ARBS but have different target populations.
  - NQF #0067, 0074, and 0070: focus on antiplatelet therapy, LDL control, and beta blocker therapy for CAD patients.

Pre-meeting public and member comments

-

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0066

Measure Title: Coronary Artery Disease: ACE Inhibitor or ARB Therapy-Diabetes or Left Ventricular Systolic Dysfunction (LVEF <40%)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: 4/29/2016

Instructions

- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- Respond to all questions as instructed with answers immediately following the question. All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 10 pages (*incudes questions/instructions*; minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- Health outcome: [3] a rationale supports the relationship of the health outcome to processes or structures of care. Applies to patient-reported outcomes (PRO), including health-related quality of life/functional status, symptom/symptom burden, experience with care, health-related behavior.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- Efficiency: [6] evidence not required for the resource use component.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the U.S. Preventive Services Task Force (USPSTF) grading definitions and methods, or Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.

5. 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 multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)
Outcome

☐ Health outcome: Click here to name the health outcome
☐Patient-reported outcome (PRO): Click here to name the PRO
PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors
☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome
☒ Process: Prescription of ACE inhibitor or ARB for patients with CAD and diabetes or LVSD
☐ Structure: Click here to name the structure
☐ Other: Click here to name what is being measured

HEALTH OUTCOME/PRO PERFORMANCE MEASURE *If not a health outcome or PRO, skip to 1a.3*

1a.2. Briefly state or diagram the path between the health outcome (or PRO) and the healthcare structures, processes, interventions, or services that influence it.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process, intervention, or service (*i.e., influence on outcome/PRO*).

Note: For health outcome/PRO performance measures, no further information is required; however, you may provide evidence for any of the structures, processes, interventions, or service identified above.

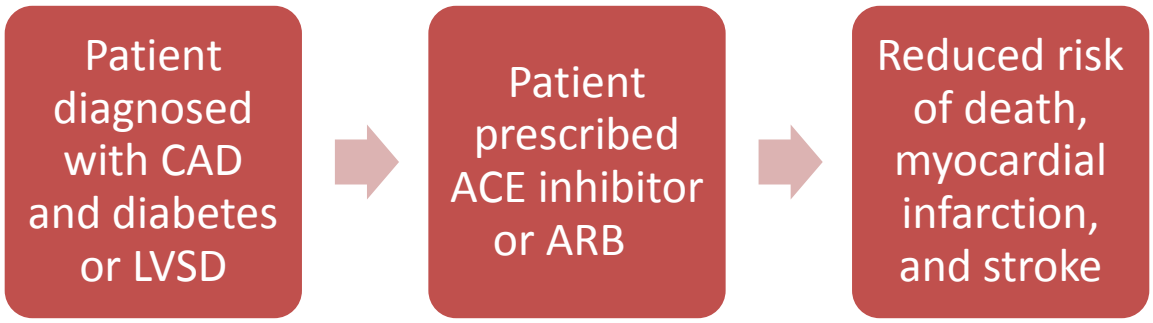
INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

1a.3. Briefly state or diagram the path between structure, process, intermediate outcome, and health outcomes. Include all the steps between the measure focus and the health outcome.

Old Submission:

Nonadherence to cardioprotective medications is prevalent among outpatients with CAD and can be associated with a broad range of adverse outcomes, including all-cause and cardiovascular mortality, cardiovascular hospitalizations, and the need for revascularization procedures.

In the absence of contraindications, ACE inhibitors or ARBs are recommended for all patients with a diagnosis of CAD and diabetes or reduced left ventricular systolic function23. ACE inhibitors remain the first choice, but ARBs can now be considered a reasonable alternative. Both pharmacologic agents have been shown to decrease the risk of death, myocardial infarction, and stroke. Additional benefits of ACE inhibitors include the reduction of diabetic symptoms and complications for patients with diabetes.



1a.3.1. What is the source of the systematic review of the body of evidence that supports the performance measure?
☒ Clinical Practice Guideline recommendation – ***complete sections 1a.4, and 1a.7***
☐ US Preventive Services Task Force Recommendation – ***complete sections 1a.5 and 1a.7***
☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*) – ***complete sections 1a.6 and 1a.7***
☐ Other – ***complete section 1a.8***

Please complete the sections indicated above for the source of evidence. You may skip the sections that do not apply.

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION

1a.4.1. Guideline citation (*including date*) and **URL for guideline** (*if available online*):

Old Submission:

Fraker JD, Fihn SD, writing on behalf of the 2002 Chronic Stable Angina Writing Committee. 2007 chronic angina focused update of the ACC/AHA 2002 Guidelines for the Management of Patients with Chronic Stable Angina: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines Writing Group to Develop the Focused Update of the 2002 Guidelines for the Management of Patients with Chronic Stable Angina. J Am Coll Cardiol. 2007;50:2264-2274.

Fihn SD, Gardin JM, Abrams J, Berra K, Blankenship JC, Dallas P, Douglas PS, Foody JM, Gerber TC, Hinderliter AL, King SB III, Kligfield PD, Krumholz HM, Kwong RYK, Lim MJ, Linderbaum JA, Mack MJ, Munger MA, Prager RL, Sabik JF, Shaw LJ, Sikkema JD, Smith CR Jr, Smith SC Jr, Spertus JA, Williams SV. 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. *Circulation.* 2012;126:e354-e471

1a.4.2. Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

Old Submission:

ACE inhibitors should be started and continued indefinitely in all patients with left ventricular ejection fraction less than or equal to 40% and in those with hypertension, diabetes, or chronic kidney disease, unless contraindicated. (Class I Recommendation, Level A Evidence). (ACC/AHA, 2007)

Angiotensin receptor blockers are recommended for patients who have hypertension, have indicators for but are intolerant of ACE inhibitors, have heart failure, or have had a myocardial infarction with left ventricular ejection fraction less than or equal to 40% (Class I Recommendation, Level A Evidence). (ACC/AHA, 2007)

4.4.2.3. Renin-Angiotensin-Aldosterone Blocker Therapy (p. e405)

1. ACE inhibitors should be prescribed in all patients with SIHD who also have hypertension, diabetes mellitus, LVEF 40% or less, or CKD, unless contraindicated. (Class I Recommendation; Level A Evidence)
2. ARBs are recommended for patients with SIHD who have hypertension, diabetes mellitus, LV systolic dysfunction, or CKD and have indications for, but are intolerant of, ACE inhibitors. (Class I Recommendation, Level A Evidence)

1a.4.3. Grade assigned to the quoted recommendation with definition of the grade:

Both recommendations are Class I Recommendations with Level A Evidence.

Class I Recommendations are defined as:

Benefit >>> Risk

Procedure/Treatment **SHOULD** be performed/administered

Level A Evidence is defined as:

Multiple populations evaluated*

Data derived from multiple randomized clinical trials or meta-analyses

**Data available from clinical trials or registries about the usefulness/efficacy in different subpopulations, such as sex, age, history of diabetes, history of prior myocardial infarction, history of heart failure, and prior aspirin use*

Recommendations with a Class I, Level A designation are characterized as:

-Recommendation that procedure or treatment is useful/effective

-Sufficient evidence from multiple randomized trials or meta-analyses

1a.4.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: If separate grades for the strength of the evidence, report them in section 1a.7.*)

Old Submission:

ACC/AHA Classification of Recommendations and Levels of Evidence

Classification of Recommendations

Class I: Conditions for which there is evidence and/or general agreement that a given procedure or treatment is beneficial, useful, and effective.

Class II: Conditions for which there is conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of a procedure or treatment.

Class IIa: Weight of evidence/opinion is in favor of usefulness/efficacy.

Class IIb: Usefulness/efficacy is less well established by evidence/opinion.

Class III: Conditions for which there is evidence and/or general agreement that a procedure/treatment is not useful/effective and in some cases may be harmful.

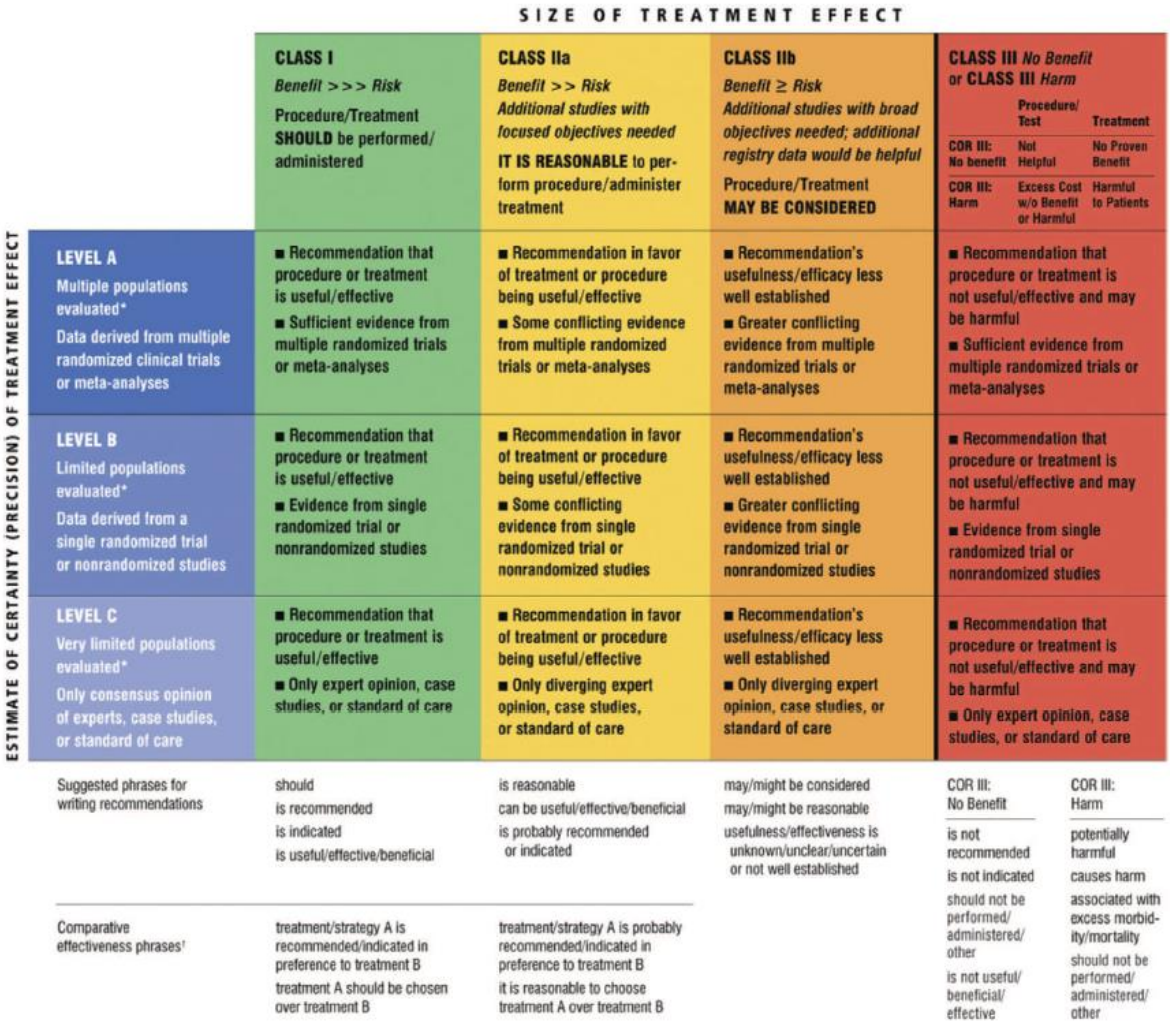
Level of Evidence

Level of Evidence A: Data derived from multiple randomized clinical trials or meta-analyses.

Level of Evidence B: Data derived from a single randomized trial, or nonrandomized studies.

Level of Evidence C: Only consensus

Current Guideline:



ESTIMATE OF CERTAINTY (PRECISION) OF TREATMENT EFFECT	SIZE OF TREATMENT EFFECT: CLASS I *Benefit >>> Risk* Procedure/Treatment **SHOULD** be performed/ administered	CLASS IIa *Benefit >> Risk* *Additional studies with focused objectives needed* **IT IS REASONABLE** to perform procedure/administer treatment	CLASS IIb *Benefit ≥ Risk* *Additional studies with broad objectives needed; additional registry data would be helpful* Procedure/Treatment **MAY BE CONSIDERED**	CLASS III *No Benefit* or CLASS III *Harm* Procedure/ Test / Treatment COR III: No benefit — Not Helpful — No Proven Benefit COR III: Harm — Excess Cost w/o Benefit or Harmful — Harmful to Patients
LEVEL A Multiple populations evaluated* Data derived from multiple randomized clinical trials or meta-analyses	■ Recommendation that procedure or treatment is useful/effective ■ Sufficient evidence from multiple randomized trials or meta-analyses	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Sufficient evidence from multiple randomized trials or meta-analyses
LEVEL B Limited populations evaluated* Data derived from a single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is useful/effective ■ Evidence from single randomized trial or nonrandomized studies	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Evidence from single randomized trial or nonrandomized studies
LEVEL C Very limited populations evaluated* Only consensus opinion of experts, case studies, or standard of care	■ Recommendation that procedure or treatment is useful/effective ■ Only expert opinion, case studies, or standard of care	■ Recommendation in favor of treatment or procedure being useful/effective ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation's usefulness/efficacy less well established ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Only expert opinion, case studies, or standard of care
Suggested phrases for writing recommendations	should is recommended is indicated is useful/effective/beneficial	is reasonable can be useful/effective/beneficial is probably recommended or indicated	may/might be considered may/might be reasonable usefulness/effectiveness is unknown/unclear/uncertain or not well established	COR III: No Benefit: is not recommended; is not indicated; should not be performed/ administered/ other; is not useful/ beneficial/ effective COR III: Harm: potentially harmful; causes harm; associated with excess morbidity/mortality; should not be performed/ administered/ other
Comparative effectiveness phrases†	treatment/strategy A is recommended/indicated in preference to treatment B treatment A should be chosen over treatment B	treatment/strategy A is probably recommended/indicated in preference to treatment B it is reasonable to choose treatment A over treatment B		

A recommendation with Level of Evidence B or C does not imply that the recommendation is weak. Many important clinical questions addressed in the guidelines do not lend themselves to clinical trials. Although randomized trials are unavailable, there may be a very clear clinical consensus that a particular test or therapy is useful or effective.

*Data available from clinical trials or registries about the usefulness/efficacy in different subpopulations, such as sex, age, history of diabetes, history of prior myocardial infarction, history of heart failure, and prior aspirin use.

†For comparative effectiveness recommendations (Class I and IIa; Level of Evidence A and B only), studies that support the use of comparator verbs should involve direct comparisons of the treatments or strategies being evaluated.

1a.4.5. Citation and URL for methodology for grading recommendations (*if different from 1a.4.1*)**:**

Not Applicable

1a.4.6. If guideline is evidence-based (rather than expert opinion), are the details of the quantity, quality, and consistency of the body of evidence available (e.g., evidence tables)?

☒ Yes → ***complete section 1a.7***

☐ No → ***report on another systematic review of the evidence in sections 1a.6 and 1a.7; if another review does not exist, provide what is known from the guideline review of evidence in 1a.7***

1a.5. UNITED STATES PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

1a.5.1. Recommendation citation (*including date*) and **URL for recommendation** (*if available online*):

1a.5.2. Identify recommendation number and/or page number and **quote verbatim, the specific recommendation**.

1a.5.3. Grade assigned to the quoted recommendation with definition of the grade:

1a.5.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: the grading system for the evidence should be reported in section 1a.7.*)

1a.5.5. Citation and URL for methodology for grading recommendations (*if different from 1a.5.1*)**:**

Complete section 1a.7

_____

1a.6. OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1. Citation (*including date*) and **URL** (*if available online*):

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

Complete section 1a.7

_____

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

If more than one systematic review of the evidence is identified above, you may choose to summarize the one (or more) for which the best information is available to provide a summary of the quantity, quality, and consistency of the body of evidence. Be sure to identify which review is the basis of the responses in this section and if more than one, provide a separate response for each review.

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

Benefits of ACE inhibitor/ARB therapy for patients with ischemic heart disease.

1a.7.2. Grade assigned for the quality of the quoted evidence with definition of the grade:

Both recommendations were assigned Level A evidence.

Level A evidence is defined as:

Multiple populations evaluated*

Data derived from multiple randomized clinical trials or meta-analyses

**Data available from clinical trials or registries about the usefulness/efficacy in different subpopulations, such as sex, age, history of diabetes, history of prior myocardial infarction, history of heart failure, and prior aspirin use*

1a.7.3. Provide all other grades and associated definitions for strength of the evidence in the grading system.

See grid provided in question 1a.4.4. for definitions of evidence grades

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range: 1995-2011

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

The cited body of evidence includes 6 randomized controlled trials and 2 meta-analyses

1a.7.6. What is the overall quality of evidence <u>across studies</u> in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

The quality of evidence and associated certainty are strong and this prompted the ACC/AHA Guideline 1A recommendation. The guideline states "clinical studies have demonstrated significant reductions in the incidence of acute myocardial infarction, unstable angina, and the need for coronary revascularization in patients after myocardial infarction with left ventricular dysfunction, independent of etiology" associated with the use of ACE inhibitors. The guideline further states that ARBs "significantly reduce LV mass and stroke incidences". Both drugs are associated with significant benefits for patients with ischemic heart disease.

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) <u>across studies</u> in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

ACE inhibitor-based regimens were associated with a 19% reduction in risk for stroke, a 16% reduction in risk for ischemic heart disease, and a 27% reduction in the risk for heart failure for each 5-mm Hg reduction in blood pressure. ARB-based regimens were associated with a 26% reduction in risk for stroke, 17% reduction in risk for ischemic heart disease, and a 12% reduction in risk for heart failure for each 5-mm Hg reduction in blood pressure.

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

The guideline does not mention any specific harms that were studied as part of the body of evidence. However, in their classification of the recommendations they assigned both as Class I recommendations with Level A evidence which indicates that anticipated benefits far outweigh potential harms. Additionally, there are no class III (harm) recommendations associated with the use of ACE inhibitors or ARBs in the guideline.

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for <u>each</u> new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

1. Rosendorff C, Lackland DT, Allison M, Aronow WS, Black HR, Blumenthal RS, Cannon CP, de Lemos JA, Elliott WJ, Findeiss L, Gersh BJ, Gore JM, Levy D, Long JB, O'Connor CM, O'Gara PT, Ogedegbe O, Oparil S, White WB; on behalf of the American Heart Association, American College of Cardiology, and American Society of Hypertension. Treatment of hypertension in patients with coronary artery disease: a scientific statement from the American Heart Association, American College of Cardiology, and American Society of Hypertension. *Hypertension.* 2015;65:1-36.
2. This is a scientific statement that provides recommendations regarding the treatment and secondary prevention of hypertension, specifically in the setting of coronary artery disease.
3. The recommendations in this statement are specific to patients who have coronary artery disease and hypertension, as well as diabetes or LVSD, which represent a subset of the patients included in the measure. The recommendations are consistent with those cited above in support of the measure.
4. The conclusions and recommendations put forth in this scientific statement are consistent with those in the 2012 stable ischemic heart disease guideline cited above.

1a.8 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.8.1 What process was used to identify the evidence?

1a.8.2. Provide the citation and summary for each piece of evidence.

1. Evidence, Performance Gap, Priority – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all subcriteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form
0066_Evidence_Attachment.docx

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- disparities in care across population groups.

1b.1. **Briefly explain the rationale for this measure** *(e.g., the benefits or improvements in quality envisioned by use of this measure)*
In the absence of contraindications, ACE inhibitors or ARBs are recommended for all patients with a diagnosis of coronary artery disease and diabetes or reduced left ventricular systolic function. ACE inhibitors remain the first choice, but ARBs can now be considered a reasonable alternative. Both pharmacologic agents have been shown to decrease the risk of death, myocardial infarction, and stroke. Additional benefits of ACE inhibitors include slowed disease progression and reduction of complications for patients with diabetes.

1b.2. **Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.** *(This is required for endorsement maintenance. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included). This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*
2014 PQRS Experience Report
2014 is the most recent year for which PQRS Experience Report measure data is available. The average performance rates on Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%)
over the last several years are as follows:
2011: 63.5%
2012: 64.0%
2013: 70.0%
2014: 81.2%

It is important to note that PQRS has been and remains a voluntary reporting program. In the early years of the PQRS program, participants received an incentive for satisfactorily reporting. However, beginning in 2015, the program imposed payment penalties for non-participants based on 2013 performance.

Reference: Center for Medicare and Medicaid Services. 2014 Reporting Experience Including Trends. Available: https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/pqrs/analysisandpayment.html

1b.3. **If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.**
In a 2016 analysis of data collected in the NCDR® PINNACLE Registry® from 2009-2012 conducted by Fleming and colleagues, prescription rates of ACEI/ARBs for patients with CAD and concurrent diabetes or LVSD ranged from 69.6% to 77.6%. (1). Similarly, in a 2014 analysis of PINNACLE Registry® data from July 2008 through December 2010, Maddox et al assessed the variation in medication prescription for patients with coronary artery disease (CAD). Among eligible patients seen in outpatient cardiology practices, ACEI/ARBs were prescribed for 69.4% (55,933/80,552) at their index clinic visit. After inclusion of all visits among eligible patients occurring within the year following the index visit, the rates increased to 72.3%. Among practices, the median prescription rate of ACEI/ARBs for eligible patients at their index clinic visit was 75.5% (range 39.1-100%) and 78.1% (range 45.4-100%) after inclusion of all visits among eligible patients occurring within the year following the index visit.(2) An earlier study by Chan and colleagues analyzed 2008-9 data from the Pinnacle registry and found a similar rate (72.4%) of ACEI/ARB prescription among CAD

patients with concurrent left ventricular systolic dysfunction (LVSD) or diabetes. (3)

References:
1. Fleming LM, Jones P, Chan PS, Adin-Christian A, Maddox TM, Farmer SA. Relationship of provider and practice volume to performance measure adherence for coronary artery disease, heart failure, and atrial fibrillation: results from the National Cardiovascular Data Registry. Circ Cardiovasc Qual Outcomes. 2016;9:48-54.
2. Maddox TM, Chan PS, Spertus JA, Tang F, Jones P, Ho PM, Bradley SM, Tsai TT, Bhatt DL, Peterson PN. Variations in coronary artery disease secondary prevention prescriptions among outpatient cardiology practices: insights from the NCDR (National Cardiovascular Data Registry). J Am Coll Cardiol. 2014 Feb 18;63(6):539-46. doi: 10.1016/j.jacc.2013.09.053. Epub 2013 Oct 30.
3. Chan PS, Oetgen WJ, Buchanan D, et al. Cardiac performance measure compliance in outpatients: the American College of Cardiology and National Cardiovascular Data Registry´s PINNACLE (Practice Innovation And Clinical Excellence) program. J. Am. Coll. Cardiol. 2010; 56(1):8–14.

1b.4. **Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability.** *(This is required for endorsement maintenance. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*
While this measure is included in a federal reporting program, that program has not yet made disparities data available for us to analyze and report.

1b.5. **If no or limited data on disparities from the measure as specified is reported in 1b4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations.**
The Chan et al. article cited above includes a secondary subgroup analysis of PINNACLE data to determine whether performance varies by patient characteristics (ie, race, sex). While the authors did not find a meaningful difference in the prescription of ACEI/ARBs for patients with LVSD or diabetes based on race, they did find a marginally significant difference based on gender (72.1% for men vs 71.7% for women; adjusted RR 0.96[0.89-0.99]; p=.05] (1) A separate analysis was completed using PINNACLE data from 2009 to evaluate the impact of insurance status on the quality of CAD care. Uninsured patients were less likely to be prescribed ACEI/ARBs than privately-insured individuals (66.7% vs 75.5%, unadjusted RR=00.88; 95% CI, 0.84-0.93, P<0.001). In addition, publicly-insured patients were less likely than privately-insured patients to be prescribed ACEI/ARBs (69.1% vs 75.5%, unadjusted RR=0.91; 95% CI, 0.89-0.94; P<0.001). (2)

1. Chan PS, Oetgen WJ, Buchanan D, et al. Cardiac performance measure compliance in outpatients: the American College of Cardiology and National Cardiovascular Data Registry´s PINNACLE (Practice Innovation And Clinical Excellence) program. J. Am. Coll. Cardiol. 2010; 56(1):8–14.
2. Smolderen KG, Spertus JA, Tang F, et al. Treatment Differences by Health Insurance Among Outpatients with Coronary Artery Disease: Insights from the NCDR®. J Am Coll Cardiol. 2013 Mar 12; 61(10): 1069–1075.

1c. High Priority (previously referred to as High Impact)
The measure addresses:

- a specific national health goal/priority identified by DHHS or the National Priorities Partnership convened by NQF; OR
- a demonstrated high-priority (high-impact) aspect of healthcare (e.g., affects large numbers of patients and/or has a substantial impact for a smaller population; leading cause of morbidity/mortality; high resource use (current and/or future); severity of illness; and severity of patient/societal consequences of poor quality).

1c.1. **Demonstrated high priority aspect of healthcare**
Affects large numbers, A leading cause of morbidity/mortality, Patient/societal consequences of poor quality, Severity of illness
1c.2. **If Other:**

1c.3. **Provide epidemiologic or resource use data that demonstrates the measure addresses a high priority aspect of healthcare. List citations in 1c.4.**
Coronary heart disease (CHD) has a prevalence of approximately 6.2% among Americans over the age of 20, for a total of approximately 15.5 million adults. (1) Each year in the United States, there are an estimated 660,000 incident coronary events, 305,000 recurrent coronary events, and an additional 160,000 silent myocardial infarctions (MIs). (1) In 2013, CHD was the underlying cause of death for approximately 1 out of 7 deaths in the United States. (1) In 2011, CHD was among the top 10 most expensive discharge diagnoses, costing approximately $10.4 billion. (1) Comorbid diabetes or left ventricular systolic disorder (LVSD) are particularly serious. CAD is the main cause of death for among patients with both type 1 and type 2 diabetes (2), while diabetes

is associated with a 2 to 4 fold increase in risk of mortality associated with heart disease (3). Among patients with CAD, presence of LVSD doubles the odds of sudden cardiac death (SCD) (4).

1c.4. **Citations for data demonstrating high priority provided in 1a.3**

1. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, Das SR, de Ferranti S, Despres J-P, Fullerton HJ, Howard VJ, Huffman MD, Isasi CR, Jiménez MC, Judd SE, Kissela BM, Lichtman JH, Lisabeth LD, Liu S, Mackey RH, Magid DJ, McGuire DK, Mohler ER III, Moy CS, Muntner P, Mussolino ME, Nasir K, Neumar RW, Nichol G, Palaniappan L, Pandey DK, Reeves MJ, Rodriguez CJ, Rosamond W, Sorlie PD, Stein J, Towfighi A, Turan TN, Virani SS, Woo D, Yeh RW, Turner MB; on behalf of the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2016 update: a report from the American Heart Association. Circulation. 2016;133:e38-e360.
2. Preis SR, Pencina MJ, Hwang SJ, D'Agostino RB Sr., Savage PJ, Levy D, Fox CS. Trends in cardiovascular disease risk factors in individuals with and without diabetes mellitus in the Framingham Heart Study. Circulation.2009;120(3):212-20.
3. Berry C, Tardif JC, Bourassa MG. Coronary heart disease in patients with diabetes: part II: recent advances in coronary revascularization. J Am Coll Cardiol. 2007;49(6):643-56.
4. Reinier K, Dervan C, Singh T, et al. Increased left ventricular mass and decreased LV systolic function have independent pathways to arrythmogenesis in coronary artery disease. Heart Rhythm. 2011;8(8):1177-82.

1c.5. **If a PRO-PM (e.g. HRQoL/functional status, symptom/burden, experience with care, health-related behaviors), provide evidence that the target population values the measured PRO and finds it meaningful.** *(Describe how and from whom their input was obtained.)*

Not applicable. Not a PRO-PM.

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the subcriteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. **Subject/Topic Area** *(check all the areas that apply):*

Cardiovascular, Cardiovascular : Congestive Heart Failure, Cardiovascular : Ischemic Heart Disease, Coronary Artery Disease, Endocrine : Diabetes

De.6. **Cross Cutting Areas** *(check all the areas that apply):*

S.1. **Measure-specific Web Page** *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

The measure specifications are included as an attachment with this submission. Additional measure details may be found at https://pqrs.cms.gov/dataset/2016-PQRS-Measure-118-11-17-2015/7u66-ng8i

S.2a. **If this is an eMeasure**, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. **Data Dictionary, Code Table, or Value Sets** *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

Attachment **Attachment:** NQF0066__I9toI10_conversion.xlsx

S.3. **For endorsement maintenance**, please briefly describe any changes to the measure specifications since last endorsement date and explain the reasons.
Supporting guidelines and coding included in the measure are reviewed on an annual basis. Minor coding changes were made to the exceptions. Formatting of the measure specification details have been modified to align with current implementation of the measure.

S.4. **Numerator Statement** *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome)*
IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
Patients who were prescribed ACE inhibitor or ARB therapy

S.5. **Time Period for Data** *(What is the time period in which data will be aggregated for the measure, e.g., 12 mo, 3 years, look back to August for flu vaccination? Note if there are different time periods for the numerator and denominator.)*
Once during 12 consecutive month measurement period

S.6. **Numerator Details** *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
Numerator Definition:
Prescribed – May include prescription given to the patient for ACE inhibitor or ARB therapy at one or more visits in the measurement period OR patient already taking ACE inhibitor or ARB therapy as documented in current medication list.

FOR POPULATION 1: Patients who are 18 years and older with a diagnosis of CAD with LVEF < 40%
Report Quality Data Code G8935: Clinician prescribed angiotensin converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) therapy

FOR POPULATION 2: Patients who are 18 years and older with a diagnosis of CAD who have diabetes
Report Quality Data Code G8473: Angiotensin converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) therapy prescribed

Note: For reporting, the two populations are combined for a single reported performance score on the combined measure population. If a patient has both diabetes and LVSD, reporting criteria #2 (CAD with diabetes) will count as appropriate reporting for this patient.

S.7. **Denominator Statement** *(Brief, narrative description of the target population being measured)*
All patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12 month period who also have diabetes OR current or prior LVEF <40%

S.8. **Target Population Category** *(Check all the populations for which the measure is specified and tested if any):*
Populations at Risk, Senior Care

S.9. **Denominator Details** *(All information required to identify and calculate the target population/denominator such as definitions, specific data collection items/responses , code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
FOR POPULATION 1: Patients who are 18 years and older with a diagnosis of CAD with LVEF < 40%
Denominator Definition:
LVEF < 40% corresponds to qualitative documentation of moderate dysfunction or severe dysfunction.

Patients aged >= 18 years
AND
Diagnosis for coronary artery disease (ICD-9-CM) [reportable through 9/30/2015]: 410.00, 410.01, 410.02, 410.10, 410.11, 410.12, 410.20, 410.21, 410.22, 410.30, 410.31, 410.32, 410.40, 410.41, 410.42, 410.50, 410.51, 410.52, 410.60, 410.61, 410.62, 410.70, 410.71, 410.72, 410.80, 410.81, 410.82, 410.90, 410.91, 410.92, 411.0, 411.1, 411.81, 411.89, 412, 413.0, 413.1, 413.9, 414.00, 414.01, 414.02, 414.03, 414.04, 414.05, 414.06, 414.07, 414.2, 414.3, 414.8, 414.9, V45.81, V45.82

Diagnosis for coronary artery disease (ICD-10-CM) [reportable beginning 10/01/2015]: I20.0, I20.1, I20.8, I20.9, I21.01, I21.02, I21.09, I21.11, I21.19, I21.21, I21.29, I21.3, I21.4, I22.0, I22.1, I22.2, I22.8, I22.9, I24.0, I24.1, I24.8, I24.9, I25.10, I25.110, I25.111, I25.118, I25.119, I25.2, I25.5, I25.6, I25.700, I25.701, I25.708, I25.709, I25.710, I25.711, I25.718, I25.719, I25.720, I25.721, I25.728, I25.729, I25.730, I25.731, I25.738, I25.739, I25.750, I25.751, I25.758, I25.759, I25.760, I25.761, I25.768, I25.769, I25.790, I25.791, I25.798, I25.799, I25.810, I25.811, I25.812, I25.82, I25.83, I25.89, I25.9, Z95.1, Z95.5, Z98.61
AND
Patient encounter during the reporting period (CPT): 99201, 99202, 99203, 99204, 99205, 99212, 99213, 99214, 99215, 99241, 99242, 99243, 99244, 99245, 99304, 99305, 99306, 99307, 99308, 99309, 99310, 99324, 99325, 99326, 99327, 99328, 99334, 99335, 99336, 99337, 99341, 99342, 99343, 99344, 99345, 99347, 99348, 99349, 99350
AND
Two Denominator Eligible Visits
AND
Report Quality Data Code: G8934: Left Ventricular Ejection Fraction (LVEF) < 40% or documentation of moderately or severely depressed left ventricular systolic function

FOR POPULATION 2: Patients who are 18 years and older with a diagnosis of CAD who have diabetes
Patients aged >= 18 years
AND
Diagnosis for coronary artery disease (ICD-9-CM) [reportable through 9/30/2015]: 410.00, 410.01, 410.02, 410.10, 410.11, 410.12, 410.20, 410.21, 410.22, 410.30, 410.31, 410.32, 410.40, 410.41, 410.42, 410.50, 410.51, 410.52, 410.60, 410.61, 410.62, 410.70, 410.71, 410.72, 410.80, 410.81, 410.82, 410.90, 410.91, 410.92, 411.0, 411.1, 411.81, 411.89, 412, 413.0, 413.1, 413.9, 414.00, 414.01, 414.02, 414.03, 414.04, 414.05, 414.06, 414.07, 414.2, 414.3, 414.8, 414.9, V45.81, V45.82
Diagnosis for coronary artery disease (ICD-10-CM) [reportable beginning 10/01/2015]: I20.0, I20.1, I20.8, I20.9, I21.01, I21.02, I21.09, I21.11, I21.19, I21.21, I21.29, I21.3, I21.4, I22.0, I22.1, I22.2, I22.8, I22.9, I24.0, I24.1, I24.8, I24.9, I25.10, I25.110, I25.111, I25.118, I25.119, I25.2, I25.5, I25.6, I25.700, I25.701, I25.708, I25.709, I25.710, I25.711, I25.718, I25.719, I25.720, I25.721, I25.728, I25.729, I25.730, I25.731, I25.738, I25.739, I25.750, I25.751, I25.758, I25.759, I25.760, I25.761, I25.768, I25.769, I25.790, I25.791, I25.798, I25.799, I25.810, I25.811, I25.812, I25.82, I25.83, I25.89, I25.9, Z95.1, Z95.5, Z98.61
AND
Diagnosis for diabetes (ICD-9-CM) [reportable through 9/30/2015]: 250.00, 250.01, 250.02, 250.03, 250.10, 250.11, 250.12, 250.13, 250.20, 250.21, 250.22, 250.23, 250.30, 250.31, 250.32, 250.33, 250.40, 250.41, 250.42, 250.43, 250.50, 250.51, 250.52, 250.53, 250.60, 250.61, 250.62, 250.63, 250.70, 250.71, 250.72, 250.73, 250.80, 250.81, 250.82, 250.83, 250.90, 250.91, 250.92, 250.93
Diagnosis for diabetes (ICD-10-CM) [reportable beginning 10/01/2015]: E10.10, E10.11, E10.21, E10.22, E10.29, E10.311, E10.319, E10.321, E10.329, E10.331, E10.339, E10.341, E10.349, E10.351, E10.359, E10.36, E10.39, E10.40, E10.41, E10.42, E10.43, E10.44, E10.49, E10.51, E10.52, E10.59, E10.610, E10.618, E10.620, E10.621, E10.622, E10.628, E10.630, E10.638, E10.641, E10.649, E10.65, E10.69, E10.8, E10.9, E11.00, E11.01, E11.21, E11.22, E11.29, E11.311, E11.319, E11.321, E11.329, E11.331, E11.339, E11.341, E11.349, E11.351, E11.359, E11.36, E11.39, E11.40, E11.41, E11.42, E11.43, E11.44, E11.49, E11.51, E11.52, E11.59, E11.610, E11.618, E11.620, E11.621, E11.622, E11.628, E11.630, E11.638, E11.641, E11.649, E11.65, E11.69, E11.8, E11.9, E13.00, E13.01, E13.10, E13.11, E13.21, E13.22, E13.29, E13.311, E13.319, E13.321, E13.329, E13.331, E13.339, E13.341, E13.349, E13.351, E13.359, E13.36, E13.39, E13.40, E13.41, E13.42, E13.43, E13.44, E13.49, E13.51, E13.52, E13.59, E13.610, E13.618, E13.620, E13.621, E13.622, E13.628, E13.630, E13.638, E13.641, E13.649, E13.65, E13.69, E13.8, E13.9
AND
Patient encounter during the reporting period (CPT): 99201, 99202, 99203, 99204, 99205, 99212, 99213, 99214, 99215, 99241, 99242, 99243, 99244, 99245, 99304, 99305, 99306, 99307, 99308, 99309, 99310, 99324, 99325, 99326, 99327, 99328, 99334, 99335, 99336, 99337, 99341, 99342, 99343, 99344, 99345, 99347, 99348, 99349, 99350
AND
Two Denominator Eligible Visits
Note: For reporting, the two populations are combined for a single reported performance score on the combined measure population. If a patient has both diabetes and LVSD, reporting criteria #2 (CAD with diabetes) will count as appropriate reporting for this patient.

S.10. **Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*
Documentation of medical reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, allergy, intolerance, pregnancy, renal failure due to ACE inhibitor, diseases of the aortic or mitral valve, other medical reasons)

Documentation of patient reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, patient declined, other patient reasons)

Documentation of system reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, lack of drug availability, other reasons

attributable to the health care system)

S.11. **Denominator Exclusion Details** *(All information required to identify and calculate exclusions from the denominator such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
Exceptions are used to remove a patient from the denominator of a performance measure when the patient does not receive a therapy or service AND that therapy or service would not be appropriate due to patient-specific reasons. The patient would otherwise meet the denominator criteria. Exceptions are not absolute, and are based on clinical judgment, individual patient characteristics, or patient preferences. The ACC/AHA/PCPI exception methodology uses three categories of reasons for which a patient may be removed from the denominator of an individual measure. These measure exception categories are not uniformly relevant across all measures; for each measure, there must be a clear rationale to permit an exception for a medical, patient, or system reason. Examples are provided in the measure exception language of instances that may constitute an exception and are intended to serve as a guide to clinicians. For measure #0066: Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy-Diabetes or Left Ventricular Systolic Dysfunction (LVEF <40%), exceptions may include medical reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, allergy, intolerance, pregnancy, renal failure due to ACE inhibitor, diseases of the aortic or mitral valve, other medical reasons), patient reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, patient declined, other patient reasons), or system reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, lack of drug availability, other reasons attributable to the health care system).Although this methodology does not require the external reporting of more detailed exception data, the ACC/AHA/PCPI recommends that physicians document the specific reasons for exception in patients' medical records for purposes of optimal patient management and audit-readiness. The ACC/AHA/PCPI also advocates for the systematic review and analysis of each physician's exceptions data to identify practice patterns and opportunities for quality improvement.

Additional details are as follows:

FOR POPULATION 1: Patients who are 18 years and older with a diagnosis of CAD with LVEF < 40%
Report Quality Data Code G8936: Clinician documented that patient was not an eligible candidate for angiotensin converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) therapy (e.g., allergy, intolerance, pregnancy, renal failure due to ACE inhibitor, diseases of the aortic or mitral valve, other medical reasons) or (e.g., patient declined, other patient reasons) or (e.g., lack of drug availability, other reasons attributable to the health care system)
FOR POPULATION 2: Patients who are 18 years and older with a diagnosis of CAD who have diabetes
Report Quality Data Code G8474: Angiotensin converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) therapy not prescribed for reasons documented by the clinician (e.g., allergy, intolerance, pregnancy, renal failure due to ACE inhibitor, diseases of the aortic or mitral valve, other medical reasons) or (e.g., patient declined, other patient reasons) or (e.g., lack of drug availability, other reasons attributable to the health care system)

S.12. **Stratification Details/Variables** *(All information required to stratify the measure results including the stratification variables, definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b)*
Consistent with CMS' Measures Management System Blueprint and recent national recommendations put forth by the IOM and NQF to standardize the collection of race and ethnicity data, we encourage the results of this measure to be stratified by race, ethnicity, administrative sex, and payer.

S.13. **Risk Adjustment Type** (Select type. Provide specifications for risk stratification in S.12 and for statistical model in S.14-15)
No risk adjustment or risk stratification
If other:

S.14. **Identify the statistical risk model method and variables** *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development and testing should be addressed with measure testing under Scientific Acceptability)*
Not applicable. No risk adjustment or risk stratification

S.15. **Detailed risk model specifications** *(must be in attached data dictionary/code list Excel or csv file. Also indicate if available at measure-specific URL identified in S.1.)*
Note: Risk model details (including coefficients, equations, codes with descriptors, definitions), should be provided on a separate worksheet in the suggested format in the Excel or csv file with data dictionary/code lists at S.2b.

S.15a. **Detailed risk model specifications** (*if not provided in excel or csv file at S.2b*)
Not applicable. No risk adjustment or risk stratification

S.16. **Type of score:**
Rate/proportion
If other:

S.17. **Interpretation of Score** *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*
Better quality = Higher score

S.18. **Calculation Algorithm/Measure Logic** *(Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.)*
To calculate performance rates:
1. Find the patients who meet the initial population (ie, the general group of patients that a set of performance measures is designed to address).
2. From the patients within the initial population criteria, find the patients who qualify for the denominator (ie, the specific group of patients for inclusion in a specific performance measure based on defined criteria). Note: in some cases the initial population and denominator are identical.
3. From the patients within the denominator, find the patients who meet the numerator criteria (ie, the group of patients in the denominator for whom a process or outcome of care occurs). Validate that the number of patients in the numerator is less than or equal to the number of patients in the denominator
4. From the patients who did not meet the numerator criteria, determine if the provider has documented that the patient meets any criteria for exception when denominator exceptions have been specified [for this measure medical reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, allergy, intolerance, pregnancy, renal failure due to ACE inhibitor, diseases of the aortic or mitral valve, other medical reasons), patient reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, patient declined, other patient reasons), or system reason(s) for not prescribing ACE inhibitor or ARB therapy (eg, lack of drug availability, other reasons attributable to the health care system)]. If the patient meets any exception criteria, they should be removed from the denominator for performance calculation.

--Although the exception cases are removed from the denominator population for the performance calculation, the exception rate (ie, percentage with valid exceptions) should be calculated and reported along with performance rates to track variations in care and highlight possible areas of focus for QI.

If the patient does not meet the numerator and a valid exception is not present, this case represents a quality failure.

S.19. **Calculation Algorithm/Measure Logic Diagram URL or Attachment** *(You also may provide a diagram of the Calculation Algorithm/Measure Logic described above at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
No diagram provided

S.20. **Sampling** *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*
IF a PRO-PM, identify whether (and how) proxy responses are allowed.
Not applicable. The measure is not based on a sample.

S.21. **Survey/Patient-reported data** *(If measure is based on a survey, provide instructions for conducting the survey and guidance on minimum response rate.)*
IF a PRO-PM, specify calculation of response rates to be reported with performance measure results.
Not applicable. The measure is not based on a survey.

S.22. **Missing data** (specify how missing data are handled, e.g., imputation, delete case.)
Required for Composites and PRO-PMs.
Patient eligibility is determined by a set of defined criteria relevant to a particular measure. If data required to determine patient eligibility are missing, those patients/cases would be ineligible for inclusion in the denominator and therefore the patient/case would be deleted.

If data required to determine if a denominator eligible patient qualifies for the numerator (or has a valid exclusion/exception) are

missing, this case would represent a quality failure.

S.23. **Data Source** *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*
If other, please describe in S.24.
Electronic Clinical Data : Registry

S.24. **Data Source or Collection Instrument** *(Identify the specific data source/data collection instrument e.g. name of database, clinical registry, collection instrument, etc.)*
IF a PRO-PM, identify the specific PROM(s); and standard methods, modes, and languages of administration.
This measure is currently being used in the ACCF PINNACLE registry for the outpatient office setting

S.25. **Data Source or Collection Instrument** *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
No data collection instrument provided

S.26. **Level of Analysis** *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*
Clinician : Group/Practice, Clinician : Individual

S.27. **Care Setting** *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*
Ambulatory Care : Clinician Office/Clinic, Ambulatory Care : Urgent Care, Behavioral Health/Psychiatric : Outpatient, Home Health, Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility
If other:

S.28. **COMPOSITE Performance Measure** - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*
Not applicable. The measure is not a composite.

2a. Reliability – See attached Measure Testing Submission Form
2b. Validity – See attached Measure Testing Submission Form
DRAFT_NQF_0066_TestingAttachment_FINAL.docx

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b2-2b7)

Measure Number (*if previously endorsed*)**:** 0066
Measure Title: Chronic Stable Coronary Artery Disease: ACE Inhibitor or ARB Therapy--Diabetes or Left Ventricular Systolic Dysfunction (LVEF <40%)
Date of Submission: 4/29/2016
Type of Measure:

☐ Composite – ***STOP – use composite testing form***	☐ Outcome (*including PRO-PM*)
☐ Cost/resource	☒ Process
☐ Efficiency	☐ Structure

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- **For all measures, sections 1, 2a2, 2b2, 2b3, and 2b5 must be completed.**
- **For outcome and resource use measures**, section **2b4** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b6** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b2-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.

- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 20 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address sociodemographic variables and testing in this form refer to the release notes for version 6.6 of the Measure Testing Attachment.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **PRO-PMs and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b2. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **PRO-PMs and composite performance measures**, validity should be demonstrated for the computed performance score.

2b3. Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; [12]

AND

If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, 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). [13]

2b4. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and sociodemographic factors) that influence the measured outcome and are present at start of care; [14,15] and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b5. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** [16] **differences in performance**;

OR

there is evidence of overall less-than-optimal performance.

2b6. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b7. For **eMeasures, composites, and PRO-PMs** (or other measures susceptible to missing data), analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements 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 of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality.

12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.
13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions
15. 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 percent v. 75 percent) 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 less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE
Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure*. *Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (***must be consistent with data sources entered in S.23***)	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☐ administrative claims	☐ administrative claims
☒ clinical database/registry	☒ clinical database/registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

The data source is EHR data from the PQRS program, provided by the Center for Medicare & Medicaid Services (CMS).

1.3. What are the dates of the data used in testing?

The data are for the time period January 2014 through December 2014 and cover the entire United States.

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: **(*must be consistent with levels entered in item S.26*)**	**Measure Tested at Level of:**
☒ individual clinician	☒ individual clinician
☒ group/practice	☒ group/practice
☐ hospital/facility/agency	☐ hospital/facility/agency
☐ health plan	☐ health plan
☐ other:	☐ other:

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

The total number of physicians reporting on this measure is 2296. Of those, 1128 physicians had all the required data elements and met the minimum number of quality reporting events (10) for inclusion in the reliability analysis. For this measure, 49.1 percent of physicians are included in the analysis, and the average number of quality reporting events is 49.0 for a total of 55,272 events. The range of quality reporting events for 1128 physicians included is from 507 to 10. The average number of quality reporting events for the remaining 50.9 percent of physicians that aren't included is 0.07.

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

There were 55,272 patients included in this testing and analysis. These were the patients that were associated with physicians who had 10 or more patients eligible for this measure.

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

The same data sample was used for reliability testing and exceptions analysis.

1.8 What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example, patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).

Patient-level socio-demographic (SDS) variables were not captured as part of the testing.

2a2. RELIABILITY TESTING

***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

☐ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)

☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

Reliability of the computed measure score was measured as the ratio of signal to noise. The signal in this case is the proportion of the variability in measured performance that can be explained by real differences in physician performance. Reliability at the level of the specific physician is given by:

Reliability = Variance (physician-to-physician) / [Variance (physician-to-physician) + Variance (physician-specific-error]

Reliability is the ratio of the physician-to-physician variance divided by the sum of the physician-to-physician variance plus the error variance specific to a physician. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in physician performance.

Reliability testing was performed by using a beta-binomial model. The beta-binomial model assumes the physician performance score is a binomial random variable conditional on the physician's true value that comes from the beta distribution. The beta distribution is usually defined by two parameters, alpha and beta. Alpha and beta can be thought of as intermediate calculations to get to the needed variance estimates.

Reliability is estimated at two different points, at the minimum number of quality reporting events for the measure and at the mean number of quality reporting events per physician.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

This measure has 0.58 reliability when evaluated at the minimum level of quality reporting events and 0.87 reliability at the average number of quality events.

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Reliability at the minimum level of quality reporting events is moderate. Reliability at the average number of quality events is high.

2b2. VALIDITY TESTING

2b2.1. What level of validity testing was conducted? (*may be one or both levels*)

☐ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ **Performance measure score**

- ☐ **Empirical validity testing**
- ☒ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)

2b2.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

Face validity of the measure score as an indicator of quality was systematically assessed as follows.

After the measure was fully specified, the expert panel was asked to rate their agreement with the following statement:

The scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good and poor quality.

Scale 1-5, where 1= Strongly Disagree; 2= Disagree; 3= Neither Agree nor Disagree; 4= Agree; 5= Strongly Agree

The expert panel included 18 members. Panel members were comprised of experts from the AHA Council on Clinical Cardiology. The list of expert panel members is as follows:

Jonathan Dukes, MD
Win Shen, MD
Michelle Albert, MD, MPH
Randal Thomas, MD
Deborah L. Crabbe, MD
Paul Wang, MD
Robert L Page II, PharmD
Vera Bittner, MD
Lori Blauwet, MD
Jennifer Cook, MD
Sana Al-Khatib, MD
Jeff Washam, PharmD
Benjamin D. Levine, MD
Jose Joglar, MD
Kiran Musunuru, MD, PhD, MPH
Michael W Rich, MD
Mauricio G. Cohen, MD
Gregory Barsness, MD

To satisfy NQF's ICD-10 Conversion Requirements, we are providing the information below:

- NQF ICD-10-CM Requirement 1: Statement of intent related to ICD-10 CM
  Goal was to convert this measure to a new code set, fully consistent with the original intent of the measure.
- NQF ICD-10-CM Requirement 2: Coding Table
  See attachment in S.2b
- NQF ICD-10-CM Requirement 3: Description of the process used to identify ICD-10 codes
  The PCPI's ICD-10 conversion approach was used to identify ICD-10 codes for this measure. The PCPI uses the General Equivalence Mappings (GEMs) as a first step in the identification of ICD-10 codes. We then review the ICD-10 codes to confirm their inclusion in the measure is consistent with the measure intent, making additions or deletions as needed. We have two RHIA-credentialed professionals on our staff who review all ICD-10 coding. For measures included in PQRS, the ICD-10 codes have also been reviewed and vetted by the CMS contractor. Comments received from stakeholders related to ICD-10 coding are first reviewed internally. Depending on the nature of the comment received, we also engage clinical experts to advise us as to whether a change to the specifications is warranted.

2b2.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

Frequency Distribution of Ratings

1 – 1 responses (Strongly Disagree)
2 – 0 responses (Disagree)
3 – 0 responses (Neither Agree nor Disagree)
4 – 7 responses (Agree)
5 – 10 responses (Strongly Agree)

The results of the expert panel rating of the validity statement were as follows: N = 17; Mean rating = 4.39 and 94.4% of respondents either agree or strongly agree that this measure can accurately distinguish good and poor quality.

2b2.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Based on the mean rating by the expert panel, this measure is valid as specified.

2b3. EXCLUSIONS ANALYSIS
NA ☐ no exclusions — *skip to section 2b4*

2b3.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

Exceptions include:

- Documentation of medical reason(s) for not prescribing angiotensin converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) therapy.
- Documentation of patient reason(s) for not prescribing angiotensin converting enzyme (ACE) inhibitor or angiotensin receptor blocker (ARB) therapy.

Exceptions were analyzed for frequency across providers.

2b3.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

Exceptions Analysis:

Amongst the 1,128 physicians with the minimum (10) number of quality reporting events, there were a total of 2,222 exceptions reported. The average number of exceptions per physician in this sample is 2.0. The overall exception rate is 3.9%.

2b3.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

Exceptions are used to remove a patient from the denominator of a performance measure when the patient does not receive a therapy or service AND that therapy or service would not be appropriate due to patient-specific reasons.

Without these being removed, the performance rate would not accurately reflect the true performance of each physician, which would result in an increase in performance failures and false negatives.

AHA recommends that physicians document the specific reasons for exception in patients' medical records for purposes of optimal patient management and audit-readiness. AHA also advocates for the systematic review and analysis of each physician's exceptions data to identify practice patterns and opportunities for quality improvement.

2b4. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b4.1. What method of controlling for differences in case mix is used?
☒ **No risk adjustment or stratification**
☐ **Statistical risk model with** Click here to enter number of factors **risk factors**
☐ **Stratification by** Click here to enter number of categories **risk categories**
☐ **Other,** Click here to enter description

2b4.2. If an outcome or resource use measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

Not applicable

2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or sociodemographic factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*)

Not applicable

2b4.4a. What were the statistical results of the analyses used to select risk factors?

Not applicable

2b4.4b. Describe the analyses and interpretation resulting in the decision to select SDS factors (e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects)

Not applicable

2b4.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)

Not applicable

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
If stratified, skip to 2b4.9

2b4.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

Not applicable

2b4.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

Not applicable

2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

Not applicable

2b4.9. Results of Risk Stratification Analysis:

Not applicable

2b4.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)

Not applicable

2b4.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

Not applicable

2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b5.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

Measures of central tendency, variability, and dispersion were calculated.

2b5.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

Based on the sample of 1,128 included physicians, the mean performance rate is 0.71 the median performance rate is 0.74 and the mode is 0.67. The standard deviation is 0.19. The range of the performance rate is 1.00, with a minimum rate of 0.00 and a maximum rate of 1.00. The interquartile range is 0.18 (0.64-0.83).

2b5.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

The range of performance from 0.64 to 0.83 suggests there's clinically meaningful variation across physicians' performance.

2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

Note: *This item is directed to measures that are risk-adjusted (with or without SDS factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without SDS factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b6.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

This test was not performed for this measure.

2b6.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

This test was not performed for this measure.

2b6.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

This test was not performed for this measure.

2b7. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b7.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

Data are not available to complete this testing.

2b7.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

Data are not available to complete this testing.

2b7.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

Data are not available to complete this testing.

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition,
Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)
If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. **To what extent are the specified data elements available electronically in defined fields?** *(i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields)*

ALL data elements are in defined fields in electronic clinical data (e.g., clinical registry, nursing home MDS, home health OASIS)

3b.2. **If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources.**

3b.3. **If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL.**

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.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 and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.**

IF a PRO-PM, consider implications for both individuals providing PROM data (patients, service recipients, respondents) and those whose performance is being measured.

We have not identified any areas of concern or made any modifications as a result of testing and operational use of the measure in relation to data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, and other feasibility issues unless otherwise noted.

3c.2. **Describe any fees, licensing, or other requirements to use any aspect of the measure as specified** *(e.g., value/code set, risk model, programming code, algorithm)*.

The Measures, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, eg, use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the Measures for commercial gain, or incorporation of the Measures into a product or service that is sold, licensed or distributed for commercial gain. Commercial uses of the Measures require a license agreement between the user and the AMA, (on behalf of the PCPI), ACC or AHA. Neither the AMA, ACC, AHA, PCPI nor its members shall be responsible for any use of the Measures.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. **Current and Planned Use**

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Planned	Current Use (for current use provide URL)
Payment Program	Public Reporting Physician Quality Reporting System (PQRS) http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/MeasuresCodes.html Quality Improvement with Benchmarking (external benchmarking to multiple organizations) NCDR Pinnacle Registry http://cvquality.acc.org/en/NCDR-Home/Registries/Outpatient-Registries.aspx

4a.1. **For each CURRENT use, checked above, provide:**

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included

1.Physician Quality Reporting System (PQRS)-Sponsored by the Centers for Medicare and Medicaid Services (CMS)
Purpose: PQRS is a national reporting program that uses a combination of incentive payments and payment adjustments to promote reporting of quality information by eligible professionals (EPs). The program provides an incentive payment to practices with EPs (identified on claims by their individual National Provider Identifier [NPI] and Tax Identification Number [TIN]). Eps satisfactorily report data on quality measures for covered Physician Fee Schedule (PFS) services furnished to Medicare Part B Fee-for-Service (FFS) beneficiaries (including Railroad Retirement Board and Medicare Secondary Payer). Beginning in 2015, the program also applies a payment adjustment to EPs who do not satisfactorily report data on quality measures for covered professional services in 2013.
Source: http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/index.html CMS has implemented a phased approach to public reporting performance information on the Physician Compare Web site. CMS announced through rulemaking their plans to make all PQRS individual EP level PQRS measures available for public reporting annually, including making the 2016 PQRS individual EP level data available for public reporting on Physician Compare in late 2017.

2. PINNACLE Registry (URL: http://cvquality.acc.org/en/NCDR-Home/Registries/Outpatient-Registries.aspx)
The PINNACLE Registry® is cardiology´s largest outpatient quality improvement registry, capturing data on coronary artery disease, hypertension, heart failure and atrial fibrillation.The PINNACLE Registry® continues to grow rapidly, with more than 5,700 providers representing over 1,500 unique office locations across the U.S submitting data to the registry. As of the fourth quarter of 2014, the registry has more than 28 million patient encounter records. PINNACLE assists practices in understanding and improving care through the production and distribution of quarterly performance reports. These reports, covering all valid patient encounters, detail adherence to 28 cardiovascular clinical measures at the physician, location, and practice levels across coronary artery disease, hypertension, heart failure and atrial fibrillation. All jointly developed ACC/AHA/PCPI performance measures for these topics are reported by the registry.

4a.2. **If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons?** (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)
Not applicable.

4a.3. **If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement.** (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)
Not applicable.

4b. Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b.1. **Progress on Improvement. (Not required for initial endorsement unless available.)**
Performance results on this measure (current and over time) should be provided in 1b.2 and 1b.4. Discuss:

- **Progress (trends in performance results, number and percentage of people receiving high-quality healthcare)**
- **Geographic area and number and percentage of accountable entities and patients included**

2014 PQRS Experience Report
2014 is the most recent year for which PQRS Experience Report measure data is available. The average performance rates on Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%)
over the last several years are as follows:
2011: 63.5%
2012: 64.0%
2013: 70.0%
2014: 81.2%

It is important to note that PQRS has been and remains a voluntary reporting program. In the early years of the PQRS program, participants received an incentive for satisfactorily reporting. However, beginning in 2015, the program imposed payment penalties for non-participants based on 2013 performance.

Reference: Center for Medicare and Medicaid Services. 2014 Reporting Experience Including Trends. Available: https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/pqrs/analysisandpayment.html

4b.2. **If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.**
While the ACC/AHA/PCPI creates measures with an ultimate goal of improving the quality of care, measurement is a mechanism to drive improvement but does not equate with improvement. Measurement can help identify opportunities for improvement with actual improvement requiring making changes to health care processes and structure. In order to promote improvement, quality measurement systems need to provide feedback to front-line clinical staff in as close to real time as possible and at the point of care whenever possible. (1)

1. Conway PH, Mostashari F, Clancy C. The future of quality measurement for improvement and accountability. JAMA. 2013 Jun 5;309(21):2215-6.

4c. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4c.1. **Were any unintended negative consequences to individuals or populations identified during testing; OR has evidence of unintended negative consequences to individuals or populations been reported since implementation? If so, identify the negative unintended consequences and describe how benefits outweigh them or actions taken to mitigate them.**
We are not aware of any unintended consequences at this time, but we take unintended consequences very seriously and therefore continuously monitor to identify actions that can be taken to mitigate them.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures
Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.
Yes

5.1a. **List of related or competing measures (selected from NQF-endorsed measures)**
0067 : Chronic Stable Coronary Artery Disease: Antiplatelet Therapy
0070 : Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)
0074 : Chronic Stable Coronary Artery Disease: Lipid Control
0081 : Heart Failure (HF): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy for Left Ventricular Systolic Dysfunction (LVSD)
1522 : ACE/ARB Therapy at Discharge for ICD implant patients with Left Ventricular Systolic Dysfunction
1662 : Angiotensin Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy
2467 : Adherence to ACEIs/ARBs for Individuals with Diabetes Mellitus

5.1b. **If related or competing measures are not NQF endorsed please indicate measure title and steward.**

5a. Harmonization

The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. **If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):**
Are the measure specifications completely harmonized?
No

5a.2. **If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.**
While this measure's specifications are harmonized with existing measures where possible, there are several key differences between this measure and other existing related measures. The first group of related measures (NQF #1662, 1522, 0081, 2467)all have a similar focus on the prescription of ACEI/ARBs. However they all have different target populations, with measure #1662 focusing on patients with chronic kidney disease (CKD), measure #1522 being a facility-level measure focusing on patients with an ICD implant, measure #0081 focusing on patients with a diagnosis of heart failure and left ventricular ejection fraction <40%, and measure #2467 focusing on medication adherence among patients with diabetes. This group of measures reflect the importance of ACEI/ARBs among a variety of patient populations, that are distinct from the patient population included in this measure. We believe that the measures are complementary rather than competing, and differences in the measure specifications are a result of the differences in the target patient population. These differences should not result in any additional data collection burden. The second group of related measures (NQF #0067, 0074, and 0070) all focus on different aspects of care for patients with CAD. Measure #0067 focuses on use of antiplatelet therapy, while measure #0074 focuses on LDL control, and measure #0070 focuses on the use of beta-blocker therapy. We view these measures as complementary measures that, when taken together, provide a rounded view of the quality of care for patients with CAD. While these measures share a focus on the patient population with CAD, differences in measure specifications are reflective of the different care processes being targeted in each measure. We don't believe that these differences result in any additional data collection burden.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. **If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):**
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)
This measure addresses a distinct target population and/or quality action from other related measures, as described above. The measures are complementary to form a well-rounded view of the quality of care for patients with CAD.

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment:

Contact Information

Co.1 **Measure Steward (Intellectual Property Owner):** American Heart Association
Co.2 **Point of Contact:** Melanie, Shahriary, melanie.shahriary@heart.org, 301-651-7548-
Co.3 **Measure Developer if different from Measure Steward:** American Heart Association
Co.4 **Point of Contact:** Melanie, Shahriary, melanie.shahriary@heart.org, 301-651-7548-

Additional Information

Ad.1 **Workgroup/Expert Panel involved in measure development**
Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
Work Group members:
Joseph Drozda, MD, FACC (Co-Chair) (cardiology; methodology)
Joseph V. Messer, MD, MACC, FAHA (Co-Chair) (cardiology)
John Spertus, MD, FACC, FAHA (Co-Chair) (cardiology)
Bruce Abramowitz, MD, FACC (interventional cardiology; measure implementation)
Karen Alexander, MD, FACC (cardiology; geriatrics)
Craig T. Beam, CRE (patient representative)
Robert O. Bonow, MD, MACC, FAHA, FACP (cardiology)
Jill S. Burkiewicz, PharmD, BCPS (pharmacy)
Michael Crouch, MD, MSPH (family medicine)
David C. Goff, Jr., MD, PhD, FAHA, FACP (internal medicine)
Richard Hellman, MD, FACP, FACE (endocrinology)
Thomas James, III, FACP, FAAP (health plan representative)
Marjorie L. King, MD, FACC, MAACVPR (cardiology; cardiac rehabilitation)
Edison A. Machado, Jr., MD, MBA (measure implementation)
Eduardo Ortiz, MD, MPH (internal medicine; guideline development)
Michael O'Toole, MD, FACC (cardiology; electrophysiology; measure implementation)
Stephen D. Persell, MD, MPH (internal medicine; measure implementation)
Jesse M. Pines, MD, MBA, MSCE, FAAEM (emergency medicine)
Frank J. Rybicki, MD, PhD (radiology)
Lawrence B. Sadwin (patient representative)
Joanna D. Sikkema, MSN, ANP-BC, FAHA (cardiology)
Peter K. Smith, MD (thoracic surgery)
Patrick J. Torcson, MD, FACP, MMM (hospital medicine)
John B. Wong MD, FACP (internal medicine)

ACCF, AHA, and PCPI measures are developed through cross-specialty, multi-disciplinary work groups. All medical specialties and other health care professional disciplines participating in patient care for the clinical condition or topic under study are invited to participate as equal contributors to the measure development process. In addition, the ACCF, AHA and PCPI strive to include on their work groups individuals representing the perspectives of patients, consumers, private health plans, and employers. This broad-based approach to measure development ensures buy-in on the measures from all stakeholders and minimizes bias toward any individual specialty or stakeholder group. All work groups have at least two co-chairs who have relevant clinical and/or measure development expertise and who are responsible for ensuring that consensus is achieved and that all perspectives are voiced.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 **Year the measure was first released:** 2003
Ad.3 **Month and Year of most recent revision:** 09, 2015
Ad.4 **What is your frequency for review/update of this measure?** Supporting guidelines, specifications, and coding for this measure are reviewed annually
Ad.5 **When is the next scheduled review/update for this measure?** 09, 2016

Ad.7 **Disclaimers:** Physician Performance Measures (Measures) and related data specifications have been developed by the American Medical Association (AMA) - convened Physician Consortium for Performance Improvement(R) (PCPI[R]), American College of Cardiology (ACC) and American Heart Association (AHA). These Measures are not clinical guidelines and do not establish a standard of medical care, and have not been tested for all potential applications. The Measures, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, eg, use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the Measures for commercial gain, or incorporation of the Measures into a product or service that is sold, licensed or distributed for commercial gain. Commercial uses of the Measures require a license agreement between the user and the AMA, (on behalf of the PCPI), ACC or AHA. Neither the AMA, ACC, AHA, PCPI nor its members shall be responsible for any use of the Measures.

THE MEASURES AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

Limited proprietary coding is contained in the Measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. The AMA, ACC, AHA, the PCPI and its members disclaim all liability for use or accuracy of any Current Procedural Terminology (CPT[R]) or other coding contained in the specifications.

Ad.8 **Additional Information/Comments:**

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 0076
De.2. **Measure Title:** Optimal Vascular Care
Co.1.1. **Measure Steward:** MN Community Measurement
De.3. **Brief Description of Measure:** The percentage of patients 18-75 years of age who had a diagnosis of ischemic vascular disease (IVD) and whose IVD was optimally managed during the measurement period as defined by achieving ALL of the following:
• Blood pressure less than 140/90 mmHg
• On a statin medication, unless allowed contraindications or exceptions are present
• Non-tobacco user
• On daily aspirin or anti-platelet medication, unless allowed contraindications or exceptions are present
1d.3. **Developer Rationale:** Achieving the intermediate physiological outcome targets related to blood pressure in addition to being tobacco free and use of daily aspirin and statins where appropriate are the cardiovascular patient's best mechanisms of avoiding or postponing long term complications associated with this chronic condition which affects millions of Americans. Measuring providers separately on individual targets is not as patient centric as a measure that seeks to reduce multiple risk factors for each patient. Patients with ischemic vascular disease are more likely to reduce their overall risk and maximize health outcomes by achieving several intermediate physiological targets and medication use targets.

Please note that while the all-or-none composite measure is considered to be the gold standard, reflecting best patient outcomes, the individual components may be measured as well. This is particularly helpful in quality improvement efforts to better understand where opportunities exist in moving the patients toward achieving all of the desired outcomes. Please refer to the additional numerator logic provided for each component.

S.4. **Numerator Statement:** The number of patients in the denominator whose IVD was optimally managed during the measurement period as defined by achieving ALL of the following:
• The most recent blood pressure in the measurement period has a systolic value of less than 140 mmHg AND a diastolic value of less than 90 mmHg
• On a statin medication, unless allowed contraindications or exceptions are present
• Patient is not a tobacco user
• On daily aspirin or anti-platelet medication, unless allowed contraindications or exceptions are present
S.7. **Denominator Statement:** Patients ages 18 to 75 with ischemic vascular disease who have at least two visits for this diagnosis in the last two years (established patient) with at least one visit in the last 12 months.
S.10. **Denominator Exclusions:** The following exclusions are allowed to be applied to the eligible population: permanent nursing home residents, receiving hospice or palliative care services, died or diagnosis coded in error.

De.1. **Measure Type:** Composite
S.23. **Data Source:** Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records
S.26. **Level of Analysis:** Clinician : Group/Practice

IF Endorsement Maintenance – Original Endorsement Date: Aug 10, 2009 **Most Recent Endorsement Date:** Jan 18, 2012

1d.1. **Composite Measure Construction:** all-or-none measures (e.g., all essential care processes received, or outcomes experienced, by each patient)
Component Measures (if endorsed or submitted for endorsement):

Maintenance of Endorsement -- Preliminary Analysis

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence
Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

1a. Evidence. The evidence requirements for a *process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured.

The evidence subcriterion (1a) must be met for each component of the composite – the 4 components of this composite have individual evidence forms and preliminary staff ratings.

The developer provides the following evidence for this measure: **Component #1 Blood Pressure**

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Summary of prior review in 2012:

- The developer provided a diagram illustrating the steps between the assessment of blood pressure control at each visit and reducing the risk of long term cardiovascular complications associated with hypertension.
- The developer provided two clinical guidelines to support blood pressure control:
  - ICSI Stable Coronary Artery Disease (April 2011), Address Modifiable Risk Factors and Comorbid Conditions: Comorbid conditions that could affect myocardial ischemia may include hypertension, anemia, thyroid disease, hypoxemia and others. Modifiable risk factors for coronary heart disease need to be evaluated and may include smoking, inadequate physical activity, stress, hyperlipidemia, obesity, hypertension and diabetes mellitus. Intervention involving any risk factor pertinent to the patient is encouraged and may include education, goal setting, and follow-up as necessary. **Evidence Grading: R** (Consensus statement; Consensus report; Narrative review)
  - ICSI Hypertension Diagnosis and Treatment Guideline (November 2010): The recommended target blood pressure is 140/90 mmHg or less. Based on current evidence, pursuing blood pressure goals lower than < 140/90 should be considered on an individual patient basis based on clinical judgment and patient preference. **Evidence Grading: A** (Randomized, controlled trial) ; **M** (Meta-analysis, systematic review, decision analysis, cost-effectiveness and analysis)

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**
☒ **The developer provided updated evidence for this measure:**

Updates:

- The developer provided three recommendations for blood pressure targets from the 2015 AHA/ACC/ASH Scientific Statement on the Treatment of Hypertension in Patients with Coronary Artery Disease:
  - BP Goal for patients with CAD is <140/90 mm Hg. **Class I; Level of Evidence: A**
  - The <140/90-mm Hg BP target is reasonable for the secondary prevention of cardiovascular events in

patients with hypertension and CAD. **Class IIa; Level of Evidence B**
    - A lower target BP (<130/80 mm Hg) may be appropriate in some individuals with CAD, previous MI, stroke or transient ischemic attack, or CAD risk equivalents (carotid artery disease, PAD, abdominal aortic aneurysm). **Class IIb; Level of Evidence B**
- The developer provided a systematic review of the body of the evidence supporting the treatment of hypertension for patients with cardiovascular disease to a target blood pressure goal of less than 140 systolic and less than 90 diastolic.
- The developer also provided the Quality, Quantity, and Consistency of the body of evidence which included 8 randomized control trials, 6 prospective observational studies, 1 meta-analysis including 147 RCTs and 1 meta-regression including 31 interventional trials.
- The developer noted that there was data that supported, but did not prove, a lower blood pressure target (<130/80 mm Hg) may be appropriate in some individuals with CAD.

Exception to evidence: N/A

Guidance from the Evidence Algorithm: Process measure with systematic review (Box 3)→Summary of QQC provided (Box 4) →Systematic review of recommendation for patients w/CAD concludes: Quantity: High; Quality: High; Consistency: High (Box 5a) → High

2012 discussion: *In 2012, the Committee noted that BP target values have been changing due to recent studies but seem to be <140/90 for most patients. New JNC 8 guidelines are due to be released in early 2012, at which time the developer will modify the measure specifications accordingly if needed.*

Questions for the Committee:

- *Is the Committee willing to accept the updated evidence from the 2015 AHA/ACC/ASH recommendations on the treatment of hypertension in patients with CAD?*

Preliminary rating for evidence: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

1a. Evidence. The evidence requirements for a *process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured.

The evidence subcriterion (1a) must be met for each component of the composite – the 4 components of this composite have individual evidence forms and preliminary staff ratings.

The developer provides the following evidence for this measure: **Component #2 Statin Medication**

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Summary of prior review in 2012:

- The developer provided a diagram illustrating the steps between assessing patients (age ≥21 to 75) with cardiovascular disease variables/risk to determine appropriate statin use and reducing the risk of long term cardiovascular complications associated with increased cholesterol levels.
- The developer provided two clinical guidelines to support lipid management:
    - ICSI Stable Coronary Artery Disease (April 2011), Address Modifiable Risk Factors and Comorbid Conditions: Comorbid conditions that could affect myocardial ischemia may include hypertension, anemia, thyroid disease, hypoxemia and others. Modifiable risk factors for coronary heart disease need to be evaluated and may include smoking, inadequate physical activity, stress, hyperlipidemia, obesity, hypertension and diabetes mellitus. Intervention involving any risk factor pertinent to the patient is encouraged and may include education, goal setting, and follow-up as necessary. **Evidence Grading**: **R** (Consensus statement; Consensus report; Narrative review)

- ICSI Lipid Management in Adults (October 2009): Hyperlipidemia: A fasting lipid profile should be evaluated for appropriate patients with stable coronary artery disease. Secondary prevention is important in these patients, who should be treated aggressively for hyperlipidemia. Many patients will require both pharmacologic and non-pharmacologic interventions to reach target goals. Target goals for hyperlipidemic patients with coronary artery disease include: LDL – less than 100 mg/dL; HDL – 40 mg/dL or greater; Triglycerides – less than 150 mg/dL. **Evidence Grading**: **A** (randomized controlled trial), **M** (Meta-analysis, systematic review, decision analysis, cost-effectiveness and analysis), **R** (Consensus statement; Consensus report; Narrative review)

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- The developer provided two clinical guidelines with recommendations for statin treatment:
  - ICSI Lipid Management in Adults (updated Nov 2013/completed prior to ACC/AHA release). Initiate Statin Treatment Recommendations: Clinicians should initiate statin therapy regardless of LDL, in patients with established ASCVD. **Evidence Grading**: **Strong Recommendation** (The work group is confident that the desirable effects of adhering to this recommendation outweigh the undesirable effects. This is a strong recommendation for or against. This applies to most patients.), **High Quality Evidence** (Further research is very unlikely to change our confidence in the estimate of effect)
  - 2013 ACC/AHA Guideline: Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults:
    - High-intensity statin therapy should be initiated or continued as first-line therapy in women and men <75 years of age who have clinical ASCVD*, unless contraindicated. **Class I; Level of Evidence: A**
    - In individuals with clinical ASCVD* in whom high-intensity statin therapy would otherwise be used, when high-intensity statin therapy is contraindicated or when characteristics predisposing to statin-associated adverse effects are present, moderate-intensity statin should be used as the second option if tolerated. **Class I; Level of Evidence: A**
  - The developer also provided secondary prevention recommendations from the ACC/AHA guideline for adults ≤75 years of age with clinical ASCVD who are not receiving statin therapy or receiving a low- or moderate-intensity statin. The recommendations state that moderate-intensity therapy should be used, if tolerated, when either high-intensity statin therapy is contraindicated or patient characteristics predisposing to statin associated adverse effects are present. There was not clear evidence of an additional reduction in ASCVD events from high-intensity statin therapy in patients >75.
- The developer provided a systematic review of the body of evidence supporting the prevention of secondary cardiovascular events for patients with cardiovascular disease by appropriately prescribing statin medications.
- The developer also provided the Quantity, Quality, and Consistency of the body of evidence which included 60 randomized control trials, 1 systematic review and 1 meta-analysis.

Exception to evidence: N/A

Guidance from the Evidence Algorithm: Process measure with systematic review (Box 3)→Summary of QQC provided (Box 4) →Systematic review concludes: Quantity: High; Quality: High; Consistency: High (Box 5a) → High

Questions for the Committee:

- *Is the Committee willing to accept the updated evidence for statin therapy?*

Preliminary rating for evidence: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

* Clinical ASCVD includes acute coronary syndromes, history of MI, stable or unstable angina, coronary or other arterial revascularization, stroke, TIA, or peripheral arterial disease presumed to be of atherosclerotic origin.

1a. Evidence. The evidence requirements for a health outcomes measure include providing rationale that supports the relationship of the health outcome to processes or structures of care. The guidance for evaluating the clinical evidence asks if the relationship between the measured health outcome and at least one clinical action is identified and supported by the stated rationale.

The evidence subcriterion (1a) must be met for each component of the composite – the 4 components of this composite have individual evidence forms and preliminary staff ratings.

Summary of evidence in 2012: Component #3 Tobacco Free

- The developer provided a diagram illustrating the path from assessment of tobacco status for the cardiovascular patient to tobacco-free.
- The developer provided the one clinical guideline to support tobacco status assessment and intervention:
  - ICSI Stable Coronary Artery Disease (April 2011), Address Modifiable Risk Factors and Comorbid Conditions: Comorbid conditions that could affect myocardial ischemia may include hypertension, anemia, thyroid disease, hypoxemia and others. Modifiable risk factors for coronary heart disease need to be evaluated and may include smoking, inadequate physical activity, stress, hyperlipidemia, obesity, hypertension and diabetes mellitus. Intervention involving any risk factor pertinent to the patient is encouraged and may include education, goal setting, and follow-up as necessary. Evidence Grading: R (Consensus statement; Consensus report; Narrative review)

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- The developer provided additional evidence from the United States Preventive Services Task Force (USPSTF) stating that despite considerable progress in tobacco control over the past 50 years, in 2013, an estimated 17.8% of U.S. adults and 15.9% of pregnant women aged 15 to 44 years were current cigarette smokers.
- The CDC indicated that smoking is a major cause of cardiovascular disease and that tobacco use contributes to heart disease and stroke by raising triglycerides, lowering (good) HDL cholesterol, increases clotting factors, damages cells that line blood vessels, increases the buildup of plague, and causes thickening and narrowing of blood vessels.

Guidance from the Evidence Algorithm: Health outcome measure (Box 1)→The relationship between the outcome and at least one process is identified and supported by the stated rationale→Pass

Question for the Committee:

- *Is there at least one thing that the provider can do to achieve a change in the measure results?*

Preliminary rating for evidence: ☒ **Pass** ☐ **No Pass**

1a. Evidence. The evidence requirements for a *process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured.

The evidence subcriterion (1a) must be met for each component of the composite – the 4 components of this composite have individual evidence forms and preliminary staff ratings.

The developer provides the following evidence for this measure: **Component #4 Daily Aspirin or Anti-Platelet Medication**

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Summary of prior review in 2012:

- The developer provided a diagram illustrating the steps between assessing patients with cardiovascular disease variables/risk to determine appropriate aspirin/anti-platelet use and reducing the risk of a subsequent cardiovascular event (secondary prevention).
- The developer provided two recommendations from the ICSI Stable Coronary Artery Disease (April 2011) clinical guideline:
    - Address Modifiable Risk Factors and Comorbid Conditions: Comorbid conditions that could affect myocardial ischemia may include hypertension, anemia, thyroid disease, hypoxemia and others. Modifiable risk factors for coronary heart disease need to be evaluated and may include smoking, inadequate physical activity, stress, hyperlipidemia, obesity, hypertension and diabetes mellitus. Intervention involving any risk factor pertinent to the patient is encouraged and may include education, goal setting, and follow-up as necessary. **Evidence Grading:** **R** (Consensus statement; Consensus report; Narrative review)
    - Antiplatelet Therapy: The use of one aspirin tablet daily (81-162 mg) is strongly recommended unless there are medical contraindications. The authors conclude that aspirin dose in the range of 75-150 mg should be given for the long-term prevention of serious vascular events in high risk patients, and that there may be a reduced benefit when increasing the dose over 150 mg daily. Doses available to most clinicians are in increments of 81 mg; therefore, the recommended dose range is 81-162 mg daily. **Evidence Grading:** **A** (randomized controlled trial), **R** (Consensus statement; Consensus report; Narrative review), **M** (Meta-analysis, systematic review, decision analysis, cost-effectiveness and analysis)

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- The developer provided three recommendations for antiplatelet agents/anticoagulants for patients with ischemic vascular disease from the AHA/ACCF Secondary Prevention and Risk Reduction Therapy for Patients with Coronary and Other Atherosclerotic Vascular Disease: 2011 Update:
    - Aspirin 75–162 mg daily is recommended in all patients with coronary artery disease unless contraindicated. **Class I; Level of Evidence: A**
        - Clopidogrel 75 mg daily is recommended as an alternative for patients who are intolerant of or allergic to aspirin. **Class I; Level of Evidence: B**
    - For patients with symptomatic atherosclerotic peripheral artery disease of the lower extremity, antiplatelet therapy with aspirin (75–325 mg daily) or clopidogrel (75 mg daily) should be started and continued. **Class I; Level of Evidence: A**
- The developer provided a systematic review of the body of the evidence supporting the prevention of secondary cardiovascular events for patients with cardiovascular disease by appropriately prescribing aspirin or anti-platelet medications.
- The developer also provided the Quality, Quantity, and Consistency of the body of evidence which included one meta-analysis of 22 randomized control trials; one collaborative meta-analysis involving 287 studies; 135,000 patients: therapy vs. control; and 77,000 patients comparing different anti-platelet regimens.

Exception to evidence: N/A

Guidance from the Evidence Algorithm: Process measure with systematic review (Box 3)→Summary of QQC provided (Box 4) →Systematic review concludes: Quantity: High; Quality: High; Consistency: High (Box 5a) → High

Questions for the Committee:

- *Is the Committee willing to accept the updated evidence from the AHA/ACCF?*

Preliminary rating for evidence: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities
Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- The developer provided the following trends for clinics in Minnesota for the **composite measure** from Report Year 2007-2016 (Dates of Service 2006-2015). *[Patients must meet all four component targets in the composite measure to be considered 'optimally managed'.]*

	2016***	2015**	2014	2013	2012	2011*	2010	2009	2008	2007
Mean	63.9									
Median	65.8									
Standard Deviation	0.10564 (10.6)									
Min, Max	8.6, 85.3									
Rate^	66.1	69.3	50.0	48.5	49.4	39.7	33.8	33.8	32.6	38.9
Pts. (Den)	104,395	102,654	98,803	87,345	78,886	66,910	63,241	46,779	36,126	4,662
Numerator	69,026	71,196	49,408	42,689	39,242	27,083	21,589	16,529	11,997	1,595
Total # of eligible pts.	104,494	103,006	99,550	93,761	95,482	96,270	95,791	80,907	54,708	11,740
% submit/eligible	99.9	99.7	99.2	93.1	82.6	69.5	66.0	57.8	66.0	39.7

* Blood pressure component target change based on evidence/ guidelines from < 130/80 to < 140/90
** Cholesterol management component suppressed during re-design
*** Cholesterol management component change from LDL < 100 to appropriate statin use

^The rate is a weighted average of the total population of patients for clinics submitting data

Reflects reportable clinics, patient n ≥30

- The developer provided the following trends for the **components** from Report Year 2009-2016 (Dates of Service 2008-2015):

	2016	2015	2014	2013	2012	2011	2010	2009
BP <140/90	85.0	85.2	84.9	84.1	84.0	-	-	-
ASA Use	96.7	96.6	96.6	96.5	94.7	94.2	91.9	92.5
Tobacco Free	83.0	83.5	81.4	82.9	82.6	82.7	81.2	82.4
Statin Use	95.2	-	-	-	-	-	-	-

- The developer also provided the distribution of rates among clinics, clinic rates by decile, a list of the top 10 best performers in Minnesota and other characteristics of the entities reporting data.

Disparities:

- The developer provided the following 2014 disparities data from the measure as specified:

Race	Optimal Rate*	Num/Den	Rate (2014)**	Change
White	67.2	(59,562/88,631)	50.8	16.4
Black/African-American	47.6	(1,451/3,049)	35.5	12.1
Asian	70.6	(1,098/1,554)	54.4	16.2
Multi-Racial	53.4	(234/359)	42.6	10.8
Chose not to disclose	68.5	(650/949)	54.3	14.2
American Indian/Alaska Native	51.8	(5,449/1,059)	34.6	17.2
Some Other Race	70.4	(178/253)	55.1	15.3

Native Hawaiian/Pacific Islander	71.4	(70/98)	50.0	21.4
Unknown	61.3	(19/31)	43.6	17.7
Race not Reported	62.0	(5,215/8,412)	48.9	13.1
Grand Total	66.1	(69,026/104,395)	50.0	16.1

Gender	Optimal Rate*	Num/Den	Rate (2014)**	Change
Female	62.5	(18,809/30,116)	44.8	17.7
Male	67.6%	(24,062/74,279)	52.1%	15.5
Grand Total	66.1	(69,026/104,395)	50.0	16.1

Age Band	Optimal Rate*	Num/Den	Rate (2014)**	Change
18 to 25	37.0	(20/54)	24.4	12.6
26 to 50	53.5%	(3,681/6,880)	35.9%	17.6
51 to 65	62.6	(27,902/44,563)	46.9	15.7
66 to 75	70.7%	(37,423/52,898)	54.8%	15.9
Grand Total	66.1	(69,026/104,395)	50.0	16.1

***** Cholesterol management component redesigned to appropriate statin use
** Cholesterol management component was LDL < 100

- The developer provided additional national and state-specific disparities data on cardiovascular disease and mortality.

Questions for the Committee:

- *Does the composite demonstrate that there continues to be a quality problem for patients diagnosed with IVD receiving 'optimal vascular care'?*
- *Do the individual components demonstrate overall less-than-optimal performance?*
- *Is a national performance measure still warranted?*
- *Are you aware of evidence that other disparities exist in this area of healthcare?*

Preliminary rating for opportunity for improvement: ☒ High ☐ Moderate ☐ Low ☐ Insufficient

1c. Composite - Quality Construct and Rationale
Maintenance measures – same emphasis on quality construct and rationale as for new measures.

1c. Composite Quality Construct and Rationale. The quality construct and rationale should be explicitly articulated and logical; a description of how the aggregation and weighting of the components is consistent with the quality construct and rationale also should be explicitly articulated and logical.

- This is an all-or-none composite measure. Patients must meet all four component targets in the composite measure to be considered 'optimally managed'.
  - The four components are weighted equally. The numerator is calculated at the patient level. Some of the components have an exception methodology within allowing a "free-pass" on the component if it does not apply to the patient.
- The components for this measure include:
  - Blood pressure control - Blood pressure less than 140/90 mmHg
  - Appropriate use of statins - On a statin medication, unless allowed contraindications or exceptions are present
  - Appropriate use of daily aspirin or anti-platelet medication - On daily aspirin or anti-platelet medication, unless allowed contraindications or exceptions are present
  - Tobacco free - Non-tobacco user
- The developer stated that this measure was developed to achieve multiple intermediate physiological/clinical

outcomes and medication use targets for IVD patients. By reducing modifiable risks, the overall risk of further ischemic vascular complications or additional cardiovascular events is reduced.

- The developer noted that measuring providers on individual targets is not as patient-centric as this composite measure that seeks to reduce multiple risk factors in patients with IVD and maximize health outcomes. The individual components are helpful in quality improvement efforts to better understand where opportunities exist in moving the patients toward achieving all of the desired outcomes.

Questions for the Committee:

- *Are the quality construct and a rationale for the composite explicitly stated and logical?*
- *Is the method for aggregation and weighting of the components explicitly stated and logical?*

Preliminary rating for composite quality construct and rationale:

☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1. Importance to Measure and Report

1a. Evidence to Support Measure Focus

Comments: **Blood pressure guidelines are in flux because of the SPRINT trial. Also, there are new data re: ezet plus statin. PCSK9 inhibitors are now available, but there are no outcome data available.

Reliability is high and validity is moderate.

Feasibility is moderate

Usability and use is high

**High rating for evidence for each of the components of the composite. QQC provided, with grade of A for almost all of the evidence from guidelines and systematic reviews.

**Evidence is strong in support of aspirin use, statin use, smoking cessation, and BP control in persons with known vascular disease. ACC/AHA recommendations generally give class I recommendations to all 4 of these strategies. Although JNC8 has slightly higher BP targets in general, AHA/ACC guidelines support a BP <140/90 for persons with known vascular disease. In contrast with the member comments, I believe that data is strong for any statin use in persons with vascular disease, though I agree that optimal statin use should be a high intensity statin absent any contraindications to these drugs.

**The evidence is generally very high for each sub-component of the measure with the weakest evidence being the consensus only data that was available for tobacco. The evidence applies directly to the measure with the exception of the act of prescribing a statin not necessarily being equal to the evidence of patients taking a statin. For purposes of the measure it seems a reasonable assumption to assume that the majority of patients that are prescribed a statin will in fact take it. The non-adherence would be difficult to account for and could confound any such measure to the same degree as the proportion of patients not actually taking medication. Health outcomes are supported by the relationship of the measure to the healthcare action. In terms of new studies-I do not think the SPRINT trial data has been adequately considered. ACCORD is the last actual significant trial that they reference in terms of hypertension goals. The other evidence presented is from 2015 Consensus guidelines. Also, recent recommendations for the general population for ASA should be considered with the population anticipated to benefit being aged 50-70.

1b. Performance Gap

Comments: **

**Medical group performance on this measure in Minnesota ranges from 8.6% to 85.3%

**Significant performance gap noted for composite. Significant disparities also noted, particularly in the African-American and American Indian populations. High rating for opportunity for improvement.

** Yes, ~1/3 of patients do not have optimal care per the measure. This number is higher for racial and ethnic minorities.

**The Minnesota clinical performance data demonstrates a clear gap and opportunity for improvement in Ischemic vascular disease optimal vascular care.

1c. High Priority (previously referred to as High Impact)

Comments: **Each of the 4 components has been shown to improve outcomes in patients with vascular disease. The 4 components are equally weighted because there are no compelling data to weight them otherwise.

**Composite quality construct is high with all-or-none measure and equally weighted components.

**Quality construct makes sense. Equal weighting is given to each component, which may or may not be appropriate. It is likely that

the absolute benefit of each component is not equal. In addition, it is much harder to achieve a BP <140/90 or smoking cessation than it is to prescribe aspirin or a statin.

**The component measures, their relationships to outcomes, and their additive value seem to be clearly stated, and logical. It seems to be fairly equally weighted across components.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

2a1. Reliability Specifications

Maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented.

Data source(s): electronic clinical data; electronic health record; paper medical record. This is not an eMeasure.

Specifications:

- This is a clinician-level, all-or-none composite measure.
- The numerator includes number of patients in the denominator whose IVD was optimally managed during the measurement period as defined by achieving **ALL** of the following:
  - The most recent blood pressure in the measurement period has a systolic value of less than 140 mmHg AND a diastolic value of less than 90 mmHg; and
  - On a statin medication** unless allowed contraindications or exceptions are present; and
  - Patient is not a tobacco*** user; and
  - On daily aspirin or anti-platelet medication, unless allowed contraindications or exceptions are present
- The numerator exceptions include:
  - Statin medication:
    - Exception to statin use based on very low LDL (< 40 for cardiovascular disease and < 70 for patients with diabetes)
    - Pregnancy at any time during the measurement period
    - Active liver disease (liver failure, cirrhosis, hepatitis)
    - Rhabdomyolysis
    - End stage renal disease on dialysis
    - Heart failure
    - Other provider documented reason: breastfeeding during the measurement period; woman of childbearing age not actively taking birth control during the measurement period; allergy to statin; drug interaction (valid drug-drug interactions include HIV protease inhibitors, nefazodone, cyclosporine, gemfibrozil, and danazol); intolerance (w/supporting documentation of trying a statin at least once w/in the last five years). Additionally, Myopathy and Myositis (CHOL-05) Value Set may be used to document intolerance to statins.
  - Aspirin or anti-platelet exception:
    - Prescribed anti-coagulant medication during the measurement period
    - History of gastrointestinal bleeding
    - History of intracranial bleeding
    - Bleeding disorder
    - Other provider documented reason: allergy to aspirin or anti-platelets; use of non-steroidal anti-inflammatory agents; documented risk for drug interaction; uncontrolled hypertension (SBP >180 mmHg and/or DBP >110 mmHg); gastroesophageal reflux disease (GERD)
- The denominator includes patients ages 18 to 75 with ischemic vascular disease who have at least two visits for this diagnosis in the last two years (established patient) with at least one visit in the last 12 months.

** Cholesterol component was redesigned in 2014 due to guideline changes and effective for dates of service 1/1/2015-12/31/2015.

*** E-cigarettes are not considered tobacco products.

- Eligible Specialties: Family Medicine, Internal Medicine, Geriatric Medicine, Cardiology
- Eligible Providers: Medical Doctor (MD), Doctor of Osteopathy (DO), Physician Assistant (PA), Advanced Practice Registered Nurses (APRN)
- 18-75 years of age as of January 1 of the measurement period (January 1 – December 31)
- Patients are identified as having a diagnosis of ischemic vascular disease (IVD) if they've had at least two face-to-face visits with an eligible provider in an eligible specialty with a diagnosis of IVD during the current or prior measurement period
- At least one face-to-face visit with an eligible provider in an eligible specialty for any reason during the measurement period

- The denominator exclusions include:
  - Patient was a permanent nursing home resident at any time during the measurement period
  - Patient was in hospice or receiving palliative care at any time during the measurement period
  - Patient died prior to the end of the measurement period
  - Documentation that diagnosis was coded in error
- ICD-10 codes are provided in the excel attachment titled MNCM_-0076_Optimal_Vascular_Care_Specs_Fields_RA_2-2016.xlsx
- The measure is risk-adjusted.
- The calculation algorithm is provided.
- Instructions for sampling and guidance on minimum sample size are included.
- Instructions for handling missing data are included.
- An excel template with formatted columns for data fields is provided by the developer.

Questions for the Committee :

- *Are all the data elements clearly defined? Are all appropriate codes included?*
- *Is the logic or calculation algorithm clear?*
- *Is it likely this measure is consistently implemented?*

2a2. Reliability Testing attachment
Maintenance measures – less emphasis if no new testing data provided

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers.

For maintenance measures, summarize the reliability testing from the prior review:

- Patient-level data element validity testing was conducted on 63,241 patients with IVD from 128 medical groups representing 573 clinics that submitted data to Minnesota Community Measurement for 2009 dates of service reported in 2010.
  - After data submission, in-person validation audits requiring 90% accuracy rate were conducted to compare the submission to the patient's medical record.
  - Of the 128 medical groups that submitted data in 2010, 17 groups initially failed the audit and remedy plans were developed. All 17 groups resubmitted and passed subsequent audit.

Describe any updates to testing: composite measure score testing conducted – see below

SUMMARY OF TESTING

Reliability testing level ☒ **Measure score** ☐ **Data element** ☐ **Both**

Reliability testing performed with the data source and level of analysis indicated for this measure ☒ **Yes** ☐ **No**

Method(s) of reliability testing:

- The dataset included 104,395 patients (99.9% of all eligible patients) with IVD in Minnesota and neighboring communities from 111 medical groups representing 671 clinics for dates of service from 1/1/2015 to 12/31/2015.
- The developers used a beta-binomial model to assess the signal-to-noise ratio. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in performance. The higher the reliability score, the greater is the

confidence with which one can distinguish the performance of one provider from another. This is an appropriate test for measure score reliability. A reliability of 0.70 is generally considered a minimum threshold for reliability.

Results of reliability testing:

- Reliability for the composite measure was **0.90** and **0.61** at the minimum number of patients per reportable clinic (≥30).
- The distribution of reliability scores by # of eligible patients per reportable clinic (≥30) ranged from 0.61 for 30 patients per clinic to 0.99 for 4,441 patients per clinic.

Guidance from the Reliability Algorithm: Precise specifications (Box 1)→Empirical reliability testing conducted w/measure as specified (Box 2)→Reliability testing conducted w/ composite measure score (Box 4) →Appropriate method used (Box 5) →High

Questions for the Committee:

- *Is the measure score test sample adequate to generalize for widespread implementation?*
- *Do the results from the updated testing demonstrate sufficient reliability so that differences in performance between providers can be identified?*

Preliminary rating for reliability: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

2b. Validity
Maintenance measures – less emphasis if no new testing data provided

2b1. Validity: Specifications

2b1. Validity Specifications. This section should determine if the measure specifications are consistent with the evidence.

Specifications consistent with evidence in 1a. ☒ **Yes** ☐ **Somewhat** ☐ **No**
Specification not completely consistent with evidence

Question for the Committee:

- *Are the specifications consistent with the evidence?*

2b2. Validity testing

2b2. Validity Testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

For maintenance measures, summarize the validity testing from the prior review:

- Patient-level data element validity testing conducted – see reliability testing from prior review above.
- Content and face validity was assessed through the Measurement and Reporting Committee and a panel of experts. There was consensus among the expert workgroup that the target components reflect a quality of care that will benefit patients in terms of reducing heart attack and stroke risk.

Describe any updates to validity testing: composite measure score empirical validity testing– see below

SUMMARY OF TESTING
Validity testing level ☒ **Measure score** ☐ **Data element testing against a gold standard** ☐ **Both**

Method of validity testing of the measure score:

☐ **Face validity only**
☒ **Empirical validity testing of the measure score**

Validity testing method:

- Empirical validity testing of the composite measure score was conducted by testing the correlation of a medical group's performance with their performance on the Optimal Diabetes Care measure (NQF #0729). It is expected

that the quality of care provided by a medical group to a patient with ischemic vascular disease would be of similar quality as the care provided to a patient with diabetes, therefore the respective performance measure scores should be similar. This is an appropriate method for assessing conceptually and theoretically sound hypothesized relationships.

Validity testing results:

- The linear regression analysis demonstrated an **R^2 value of 0.635** which means that 64% of the total variation in performance on the Optimal Vascular Care measure can be explained by variation in performance on the Optimal Diabetes Care measure. The other **36%** of the total variation on the Optimal Vascular Care measure remains unexplained.

Questions for the Committee:

- *Do the results demonstrate sufficient validity so that conclusions about quality can be made?*
- *Does the degree of correlation accurately reflect the quality of care provided for patients with IVD?*

2b3-2b7. Threats to Validity

2b3. Exclusions:

- The developer provided the analysis of exclusions from the 2016 Report Year (2015 Year of Service)
  - # of medical groups that submitted exclusions: 51 of 113 (45.1%)
  - Total # of exclusions: 1,249/104,395 = 1.20%
    - Patient was a permanent nursing home resident: 234
    - Patient was in hospice: 91
    - Patient deceased: 918
    - Documentation that diagnosis was coded in error: 6
- The developer stated that the upper age limit cut-off limits the frail elderly population in which the targets may not be appropriate but these exclusions capture potentially frail patients less than age 75.

Questions for the Committee:

- *Are the exclusions consistent with the evidence?*
- *Are any patients or patient groups inappropriately excluded from the measure?*
- *Are the exclusions/exceptions of sufficient frequency and variation across providers to be needed (and outweigh the data collection burden)?*

2b4. Risk adjustment: **Risk-adjustment method** ☐ **None** ☒ **Statistical model** ☐ **Stratification**

Conceptual rationale for SDS factors included ? ☐ **Yes** ☒ **No**

SDS factors included in risk model? ☐ **Yes** ☒ **No**

Risk adjustment summary:

- The developer used the following framework of criteria for each potential risk factor for adjustment:
  - Clinical/conceptual relationship and empirical association with the outcome
  - Variation in prevalence across providers
  - Present at the start of care and not represent the quality of care provided
  - Resistant to manipulation or gaming
  - Reliable and feasible data captured at a reasonable cost
  - Contribution of unique variation in the outcome
  - Potentially, improvement of the risk model; face validity and acceptability
- The developer analyzed the following risk factors for selection in the risk model:
  - Age (18-25; 26-50; 51-65)
  - Gender
  - Comorbidity (Depression)
  - Distance from Clinic
  - Insurance Product (Medicare, Medicaid, MSHO, Special Needs, Self-pay, Uninsured
- The final risk factors selected for the risk model were Age and Insurance Product
  - Gender did not show enough variation between clinics

- Depression was not selected due to the high cost of collection
- Race, ethnicity, language, and country of origin (RELO) were not considered for risk adjustment because these variables did not have a high completion rate across all clinics. The developer stated that they are continuing to work with the medical community to achieve the goal of evaluating RELO at the clinic level.

Risk Model Discrimination Statistics:

- The developer conducted an Analysis of Maximum Likelihood Estimates on the 2014 Dates of Service to compare the optimal rate of patients by insurance product (Commercial, MHCP, and Uninsured) to patients with Medicare and patient age (18-25; 26-50; 51-65) to patients aged 66-75.

Statistical Risk Model Calibration Statistics:

- The Analysis of Maximum Likelihood Estimates demonstrated that all of the results of both variables, age and insurance product, were significant, except for ages less than 26 due to the small sample size (n = 44).
- The developer also found that the only two variables that were correlated were age >65 and Medicare.

Risk Segmentation Analysis:

- The developer conducted a risk segmentation analysis from 1/1/14-12/31/14 and compared the age of the patient and insurance type. The analysis concluded that for all ages over 25 and for all insurance products, there are significant differences between categories.

Questions for the Committee:

- *Is an appropriate risk-adjustment strategy included in the measure?*
- *Are the candidate and final variables included in the risk adjustment model adequately described for the measure to be implemented?*
- *Are all of the risk adjustment variables present at the start of care? If not, describe the rationale provided.*

2b5. Meaningful difference *(can statistically significant and clinically/practically meaningful differences in performance measure scores can be identified)*:

- For each measure, the developer calculated both the individual medical group rate and the medical group average rate. High performance was identified by comparing individual medical group/clinic rates with the statewide average and categorized as 'Above', 'Average', or 'Below'.
- The developer provided the distribution of optimal vascular rates 2016 for 491 clinics.

Question for the Committee:

- *Does this measure identify meaningful differences about quality?*

2b6. Comparability of data sources/methods:

- N/A

2b7. Missing Data

- The developer evaluated over 104,395 ischemic vascular patients for 2013 Dates of Service for:
  - inappropriate measurement timeframe
  - invalid values

Variable	Inappropriate measurement timeframe	Invalid value
Blood Pressure	0.2%	0.03%
Tobacco Status documented	0.6%	0.0%
ASA or anti-platelet	3.3%	-
Statin	5.7%	-

- Patients with missing data are not excluded from the measure. Elements missing from any component are counted as a numerator component fail and remain in the denominator.

2d. Composite measure: construction

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality

construct.

- This is a patient-level all-or-none composite measure. The components are equally weighted and can be analyzed individually for the purposes of understanding opportunities for improvement within the composite.
- The developer conducted a Pearson Correlation Analysis of the individual components rates and the composite rates. The Pearson correlation coefficient value, r, ranges from +1 to -1. A value of 0 indicates that there is no association between the two variables. A value greater than 0 indicates a positive association; that is, as the value of one variable increases, so does the value of the other variable. A value less than 0 indicates a negative association; that is, as the value of one variable increases, the value of the other variable decreases.
- The developer conducted the following Pearson Correlation Analysis for each component:

Pearson Correlation Analysis		
Variable	**Mean**	**Pearson r coefficient**
Blood pressure	0.85048	0.69813
Tobacco Free	0.80901	0.71336
Daily ASA Use	0.96271	0.59223
Statin Use	0.93973	0.62327
N = 491 Optimal Care Rate = 0.63919		

- The developer concluded that practices in Minnesota demonstrate relatively high compliance for all of the components, however, there is still an opportunity for improvement at the clinic level.
- The blood pressure control and tobacco free components demonstrate the most variability, opportunity for improvement, and impact ability to achieve all four components.

Questions for the Committee:

- *Do the component measures fit the quality construct?*
- *Are the objectives of parsimony and simplicity achieved while supporting the quality construct?*

Guidance from Validity Algorithm: Specifications consistent with evidence (Box 1)→Potential threats to validity assessed (Box 2)→Empirical validity testing (Box 3) →Composite measure score validity testing conducted (Box 6) →Appropriate method for assessing conceptually and theoretically sound hypothesized relationships/Correlation of the composite measure score on this measure and other composite measure (Box 7) →High/Moderate certainty or confidence that the performance measure scores are a valid indicator of quality (Box 8a/8b) →Moderate

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2d)

2a1. & 2b1. Specifications

Comments: **All of the data elements are clearly defined. Codes are provided. It appears that this measure can be and has been consistently implemented.

**Measure specifications are clearly defined and consistent with the evidence. Measure likely to be implemented consistently. Unsure why "permanent nursing home resident" is in the denominator exclusions.

**No issues with reliability specifications

**No data elements were not clearly defined and the sampling methods and survey looks reliable and specific.

**The data on blood pressure goals have changed within the past 9 months (i.e., the SPRINT trial). New pharmaceuticals are available for the treatment of hyperlipidemia. However, the SPRINT trial is only a single trial and endpoint data are lacking for the new pharmaceutical agents (PCSK9 inhibitors).
**Specifications consistent with the evidence.
**Per my comments above, use of a high intensity statin rather than any statin would be more consistent with optimal vascular care per ACC/AHA guidelines in persons with known vascular disease.
**With statins, the evidence that we are currently basing guidelines on suggest that there is a significant difference between the intensity of statin treatment and the outcomes. With this measure the all or nothing prescribing behavior may not be consistent with the evidence for outcomes. For example, is there evidence that the lowest dose of the weakest statin if prescribed will generate clinically significant outcomes in the IVD population-not really but that would meet the outcome measure of the statin being prescribed. Measuring the LDL gives a more reliable benchmark of effect related to specific clinical outcomes.
2a2. Reliability Testing
Comments: **The dataset included 104,395 patients (99.9% of all eligible patients) with IVD in Minnesota and neighboring communities from 111 medical groups representing 671 clinics for dates of service from 1/1/2015 to 12/31/2015

- Reliability for the composite measure was 0.90 and 0.61 at the minimum number of patients per reportable clinic (≥30).
- The distribution of reliability scores by # of eligible patients per reportable clinic (≥30) ranged from 0.61 for 30 patients per clinic to 0.99 for 4,441 patients per clinic.

**Reliability testing was done through beta-binomial model measuring signal-to-noise ratio, and the results demonstrated high reliability overall. Reliability was even moderate at the minimum number of patients.
**As goal of the measure is to assess practice/clinical level performance (I believe), reliability is a function of the number of patients per practice. If target reliability is >0.7 and some practices have a reliability of 0.61 because of having only 30 total patients, the overall reliability of the measure will be weakened if there are many practices with a small number of patients. What is the distribution across practices in the number of patients?
**Reliability was tested with adequate scope. The results demonstrate sufficient reliability at the score level.
2b2. Validity Testing
Comments: **The stewards tested this measure against their Optimal Diabetes Measure and found an R-squared of 0.635. They also have community consensus that the measure is a valid indicator of care for patients with IVD.
**Construct (empirical validity testing) done through correlation testing, and results indicate moderate correlation. Pearson correlation analysis also done to support the composite approach. Again, moderate correlation demonstrated. thus overall moderate validity.
**The use of empirical validity testing with the performance on the diabetes measure strikes me as a relatively weak approach. We know that performance measures with regard to outcomes (like 30-day mortality or 30-day readmission) do not correlate strongly within hospitals (i.e. hospitals with high mortality performance for AMI don't necessarily have high mortality performance for HF or pneumonia). I'm not sure then what to make of fact that correlation in vascular measure scores with the diabetes measure scores is ~0.6 (r-squared) really means
**There were ample patients for validity testing based on what was provided. This measure should be generalizable. As a whole of individual components the measure should allow significant conclusions to be made about quality. I agree that this measure is an indicator of quality because the components of the measure have been shown with a high level of evidence to improve patient outcomes.
2b3. Exclusions Analysis
2b4. Risk Adjustment/Stratification for Outcome or Resource Use Measures
2b5. Identification of Statistically Significant & Meaningful Differences In Performance
2b6. Comparability of Performance Scores When More Than One Set of Specifications
2b7. Missing Data Analysis and Minimizing Bias
Comments: **"Exclusions are minimal.
performance varies by age and Medicare (y/n)
Risk adjustment is appropriate
Missing data are counted as failure to perform. Therefore they are not a threat to validity."
**Exclusions noted. No apparent threats with small percentage of exclusions, but still unsure regarding nursing home exclusion. Risk adjustment done through statistical risk model. Results demonstrate that Age and Insurance meet requirements as risk variables.

**Patient level data seems to indicate a relationship between race/ethnicity and measure score. Is this because of intrinsic characteristics of patients or poorer quality of care at practices that care for a larger percentage of racial and ethnic minority groups? It seems like this question needs to be explored.
**There are minimal threats to validity as the exclusions are evidence based and should be randomized across a wide sample of providers.
The differences are clinically meaningful for patients with IVD and should represent an actual difference in quality and subsequently better outcomes for those that are measured higher.
SDS variables align with the description. This is not the focus of the measure and comparisons and disparities are not part of the metric among this population. Given the potential for a sicker population a risk adjustment approach may be appropriate.
Missing data should not constitute a threat to the validity of this measure.
2d1. Composite Performance Measure – Composite Analysis
Comments ** "The developer concluded that practices in Minnesota demonstrate relatively high compliance for all of the components, however, there is still an opportunity for improvement at the clinic level. While the lowest average performance is for tobacco-free (0.809), the average optimal care rate is 0.64, indicating that there is significant opportunity to improve performance. The four components are all weighted pass/fail and failure of any one results in failure of the composite. There is no compelling evidence that this scoring system is inappropriate."
**Pearson correlation analysis also done to support the composite approach. Again, moderate correlation demonstrated. thus overall moderate validity.
**Yes, no concerns here.
**The analysis does demonstrate that the measures fit the quality construct and do add value. The aggregation and weighting rules do seem to fit the quality construct and rationale.

Criterion 3. Feasibility
Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- All data elements are in defined fields in electronic health records.
- The developer listed what they have learned over several years of operational use about data submission, providing detailed specifications, audit methods, patient confidentiality, EHR's, data collection burden, and the impact of health plans on the number of medical groups reporting this measure.
- There are no fees associated with participation and submitting data for this measure to MNCM. There are costs associated with data extraction and abstraction.

Questions for the Committee:
- *Are the required data elements routinely generated and used during care delivery?*
- *Are the required data elements available in electronic form, e.g., EHR or other electronic sources?*

Preliminary rating for feasibility: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments
Criteria 3: Feasibility

3a. Byproduct of Care Processes
3b. Electronic Sources
3c. Data Collection Strategy
Comments: **All data elements are in defined fields in electronic health records.
The developer listed what they have learned over several years of operational use about data submission, providing detailed specifications, audit methods, patient confidentiality, EHR's, data collection burden, and the impact of health plans on the number of medical groups reporting this measure.
There are no fees associated with participation and submitting data for this measure to MNCM. There are costs associated with data extraction and abstraction.
All primary care, multi-specialty clinics with cardiology services and cardiology clinics in Minnesota must report performance and

similar service providers in bordering communities report on a voluntary basis. This is 111 Medical groups representing 671 clinic sites; 2015 dates of service 104,494 patients with ischemic vascular disease."

**Data collection obtained through EHRs. No data collection barriers identified. There are costs associated with data extraction and abstraction. Moderate to high feasibility.

**Medication use, BP readings, and smoking status should be available in the medical record in structured fields in most cases. Contraindications to medications may not always be available in structured fields within EHRs

**All of the sub-measures are routinely collected in most medical practices. The way these are documented varies widely and does require some specific data entry versus allowing the EHR or a data extractor to reliably quantify all of the measures. For example, smoking status or tobacco use might be documented in a non-discrete field such as a text field but this may be difficult to reliably extract.

Criterion 4: Usability and Use
Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact /improvement and unintended consequences

4. Usability and Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No**

Accountability program details:

- Public reporting:
  - MN Community Measurement- MN HealthScores Website. Public Reporting consumer-facing website. All primary care, multi-specialty clinics with cardiology services and cardiology clinics in Minnesota (mandatory) and bordering communities (voluntary). 111 Medical groups representing 671 clinic sites; 2015 dates of service 104,494 patients with ischemic vascular disease.
  - MNCM Health Care Quality, Equity and Disparity Annual Reports. Hard-copy reports (pdf) highlighting top performers, most improved. 111 Medical groups representing 671 clinic sites; 2015 dates of service 104,494 patients with ischemic vascular disease.
- Payment:
  - MN Bridges to Excellence and MN Quality Incentive Payment System. Pay for Performance Program for top performers and those attaining improvement goals/ benchmarks. All clinics participating in MNCM's data portal who submit full population are eligible for inclusion in this program. Annual recognition event and rewards distributed.
  - Quality Incentive Payment System (QIPS). The Minnesota Quality Incentive Payment System (QIPS) is a statewide pay-for-performance system for physician clinics and hospitals. The system rewards providers for two types of accomplishment: (1) achieving absolute performance benchmarks or (2) improvements in performance over time. In 2014, MMB and DHS paid nearly $500,000 in incentive payments to providers in 259 clinics that achieved benchmarks or significantly improved care for diabetes, vascular disease, and depression.
- Regulatory and Accreditation Programs:
  - Minnesota Statewide Quality Reporting and Measurement System (SQRMS). Based on 2008 health reform state legislation; this program requires mandatory submission of data from Minnesota physician clinics that have provider specialties that are applicable to the measured population. For the Optimal Vascular Control Measure: family medicine, general practice, internal medicine, geriatric medicine and cardiology.
- Quality Improvement with Benchmarking (external benchmarking to multiple organizations):
  - Minnesota Department of Human Services (DHS). This competitive grant program is administered by the Minnesota Department of Health's Office of Minority and Multicultural Health and provides funds

to close the gap in the health status of African Americans/Africans, American Indians, Asian Americans, and Hispanics/Latinos in Minnesota as compared to Whites in the following priority health areas: Breast and cervical cancer screening, diabetes, heart disease and stroke, HIV/AIDS and sexually-transmitted diseases, immunizations, infant mortality, teen pregnancy prevention and unintentional injury and violence.

Improvement results:

- The developer provided the performance results included in Section 1b. but noted that although the state-wide average, reflecting over 104,000 patients, may appear to be making only small incremental improvements year to year, please note the increasing number of numerator cases. Between report years 2013 and 2014; prior to measure redesign, 6,719 more patients achieved optimal control targets than the previous year.

Unexpected findings (positive or negative) during implementation:

- Measure #0076 Optimal Vascular Care is currently in use; included in MN state legislation for the Statewide Measurement and Reporting System (SQRMS) for the past 6 years. In accordance with significant changes in guidelines and evidence for cholesterol management (ACC/ AHA), one component of this patient level all-or-none composite has undergone redesign. Redesigned measure construct and first year data on over 100,000 patients is included in this maintenance update.

Potential harms:

- The developer did not identify any unintended consequences during the testing, implementation and ongoing review of this measure.

***Questions for the Committee**:*

- *How can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

Preliminary rating for usability and use: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 4: Usability and Use

4a. Accountability and Transparency

4b. Improvement

4c. Unintended Consequences

Comments: **"There is a public-facing web site and a hard copy report highlights top performers and disparities.

I am not aware of unintended consequences.

**Measure is currently reported publically through MN Community Measurement HealthScores and in MNCM annual reports. Additional accountability in payment accreditation and quality improvement programs. High usability.

**This measure seems to be used in many programs in MN for public reporting and payment. Unintended consequences may arise if lower performance at practices caring for a larger proportion of racial and ethnic minority patients relates primarily to intrinsic characteristics of patients rather than the quality of care provided by the practice.

**This measure can be used for payment incentive by payers, and for quality improvement within a given practice or system. Overall it seems to be very usable and reportable. The only unintended consequences might be that if the measure is broadly applied in the interest of payment that some health care providers will be motivated to discharge, not admit, or systematically avoid the patients at highest risk and with the least control of their sub measures.

Criterion 5: Related and Competing Measures

Related or competing measures:

0067 : Chronic Stable Coronary Artery Disease: Antiplatelet Therapy
0068 : Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antiplatelet
0073 : Ischemic Vascular Disease (IVD): Blood Pressure Control
0543 : Adherence to Statin Therapy for Individuals with Cardiovascular Disease (CMS) – no longer NQF endorsed

Harmonization :

- #0068 and #0073 focus on the inpatient setting and patients discharged with AMI, CABG, or PCI.
- #0067 focuses on patients with CAD.

Pre-meeting public and member comments

Comment by: Dr. Kay E. Jewell, MD

Organization: SMT Inc.

Comments: #5655, #5660, #5661, #5662, #5663, #5664, and #5665
We have a number of concerns with the addition of the component 'on a statin medication, unless allowed contraindications or exceptions are present' to replace the LDL-C component in the NQF Measure #0076 Optimal Vascular Care: All-or-none composite measure.

This all-or-none measure as redefined with statin use as a component is misleading to the patient who will be using this measure to assess the quality of care being provided. The patient has the right to assume #0076 addresses elements of care needed to provide optimal and high quality care which will reduce their risk associated with Ischemic Vascular Disease. This assumption about the purpose of the measure is reflected in statements provided in the measure specifications, in the Developer Rationale, and in the evidence and quality section of the composite measure (1.d.2) which states "The desired goal is for the patient is to achieve multiple intermediate physiological clinical outcome and medication use targets to best reduce their overall risk of developing further ischemic vascular complications (short and long term) or an additional cardiovascular event." The patient would have the right to assume that #0076 has been appropriately designed so that the component measures reflect the major risk factors associated with ischemic vascular events AND each component measure is evidence-based AND each component measure has been found to have a positive impact on the reduction of IVD risk or events.

In this case, the simple act of prescribing a statin is being proposed as the component process measure to address the cholesterol component of ASCVD risk. This raises 3 issues of concern. a) This component measure has not been tested as a process measure and shown that it is linked to a decrease in ASCVD events. b) As it is written, it is not 'evidence-based' because it does not reflect the evidence/guideline recommendations for optimal management; rather it reflects 'minimal' management. c) There is an exclusion for those who do not tolerate statin therapy. Because of these 3 issues, the inclusion of 'statin therapy' with exception as the component measure to address cholesterol management weakens the value and potential impact of the all-or-none measure. As pointed out by Nolan and Berwick, "The all-or-none approach will amplify errors of measurement (one unreliable component measure will contaminate the whole score) so that it is essential that each of the component measures be well designed." (Nolan JAMA 2006: 1168-70)

The statin component accepting any dose of statins in high risk patients has not been proven to be effective in reducing ASCVD events or mortality. We recognize that there is a gap in care and a significant number of people with a high or moderate risk of ASCVD who are not taking statins. Some argue it's better the patients take any dose, even a low dose, than take nothing because it will still have some impact on their ASCVD risk for recurrent events. We don't know the degree to which a low dose will impact risk reduction compared to the body of evidence that supports the impact of moderate and high intensity therapy. "Potential ASCVD risk-reduction benefits. The absolute reduction in ASCVD events from moderate- or high-intensity statin therapy can be approximated by multiplying the estimated 10-year ASCVD risk by the anticipated relative-risk reduction from the intensity of statin initiated (~30% for moderate-intensity statin or ~45% for high-intensity statin therapy)."(Stone JACC 2014, p 2898) The CTT reported a significant difference between high-dose vs. moderate dose statin. They found that high dosing achieved ~ 0·5 mmol/L greater reduction in LDL-C (19 mg/dL). After 1 year, this translated into a 15% greater reduction in major vascular events and a 13% reduction in coronary death or non-fatal MI for those on more intensive treatment. (CTT Lancet 2010; 1670-1681.) More importantly, even if lower intensity will reduce ASCVD events, the 2013 ACC guidelines do not consider it 'optimal' management for the individual patient; "… individuals who merit guideline-recommended statin therapy should be

treated with the maximum appropriate intensity of a statin that does not cause adverse effects." (Stone JACC 2014, p 2912)

Accepting non-optimal dosing is a low bar for quality which may have been sufficient years ago when we were first starting with measures but we have had years of attention to LDL-C levels with the HEDIS measurement and others and even with specific LDL-C recommendations in 2002 with the ATP III Report and still patients were not been receiving optimal therapy as evidenced by the low scores in HEDIS before they retired the LDL-C measures in 2015. As a high priority public health issue, we should not be satisfied going back to a low-bar effort for an important risk factor in ASCVD risk. In addition, as an unweighted all-or-none measure, it is presumed that each of the components contribute equally to the end result, reduction in ASCVD risk. It is not clear it can be said that ordering of any statin dose in the high/very high risk people will have equal weight with BP control or smoking cessation or taking aspirin.

According to the data submitted with the measure, this component of the measure is already being achieved by 95% of those reporting, leaving little room for improvement. As a stand-alone measure, this would create a problem which is reflected in the CMS recommendation to retire the hospital measure for statins ordered at the time of discharge from the hospital after an AMI (AMI-10). (IPPS - CMS- 1655-P FedReg 04/27/2016) The fact that nearly every patient is being prescribed a statin tells us nothing about the actual level of care being provided. It cannot be assumed that clinicians have ordered the appropriate intensity of statin to match the patient's level of ASCVD risk or that they will monitor and titrate treatment to the desired response or that they will follow the recommendations in the guidelines published by ACC/AHA, NLA (Jacobson, J Clin Lipidol. 2015: 129-169; 2015: S1-S122) , or the Decision Pathway on non-statins. (Lloyd-Jones JACC, 2016: doi:10.1016/j.jacc.2016.03.519) There are a number of published studies which show this does not occur. Experience with hospital measures and statins ordered at discharge after an AMI: at 12 months, the studies showed that only 26% had been adjusted to reach target. The dose at the time of discharge remained the same after 1 yr. (Arnold JACC 2013: Arnold Circ 2014: 1303-1309; Werner JAMA 2006: 2694-702; Fonarow JAMA 2007: 61-70; Glickman Am Ht J 2007: 1206-20.) Maddox et al reported that not even cardiologists prescribed at a dose necessary to achieve LDL-C goal when that was the recommendation (Maddox 2014). Schoen et al reported on their experience, that only 9.4% of their patients in an academic setting would be considered to be on the dose recommended. 43% would need to have an increase in their dose. (Schoen Am J Med 2014). This has serious implications based on the number of patients reported in this measure and impacted by the recommendations for high-intensity statins and for those at high risk. The measure should be designed to address this concern and to promote the recommended standards of 'optimal care'.

There are two types of process measures for cholesterol management that would come closer to demonstrating 'optimal care': a process measure that requires that the prescribed statin match the intensity of therapy with level of risk, e.g. high intensity with high risk and/or a measure that reports whether the prescription was filled. The challenge is that even well-defined process measures have not been shown to be as effective in improving the quality of care as reflected in patient outcome. Even in this composite measure, two of the components are outcome measures – BP control and nonsmoking. Optimal care for cholesterol management would best be demonstrated with an outcome/intermediate outcome measure of actual 'optimal control'.

For cholesterol management, optimal care would demonstrate that the patient has achieved the desired therapeutic response as reflected by the percent LDL-C reduction or a threshold LDL-C level achieved (Table 3, C.1.d) and further addressed in the ACC Decision Pathway for non-statin therapy.(Lloyd-Jones 2016) We recognize that the 2013 guidelines did not recommend treating to a specific LDL-C level as a 'fixed target' but their concern was related more to undertreatment of those at high risk: "If a moderate- or low-intensity statin results in an LDL-C level <100 mg/dL in a patient with ASCVD, the evidence suggests that a high-intensity statin, if tolerated, provides a greater reduction in ASCVD events." (Stone JACC 2014 p.2913,). The guideline does recognize the LDL-C as an indicator of response; the LDL-C level does have clinical meaning, as stated repeatedly in the guideline text and in the Evidence Statements (Appendix 4):
"Classifying specific statins and doses by the percent reduction in LDL-C level is based on evidence that the relative reduction in ASCVD risk from statin therapy is related to the degree by which LDL-C-C is lowered."

"…extensive evidence shows that each 39-mg/dL reduction in LDL-C by statin therapy reduces ASCVD risk by about 20%."

It is difficult to avoid the fact that for some conditions, there is a recommended threshold that signals optimal care that is directly linked to improved clinical outcome for the patient, regardless of how the patient or the physician accept the challenge of achieving that threshold. Examples are blood pressure, LDL-C, glucose control and weight/obesity. The bottom line is that the evidence supports a link between the reduction in LDL-C including threshold levels and the reduction in ASCVD events; it is helpful to the patient and the physician to have a general number which influences the decision about the adequacy of the therapeutic response.

In order to avoid an unintended consequence of creating a new gap in care for this high risk population, we should reconsider excluding those who have contraindications or exclusions from this component of the all-or-none measure. This measure applies to those who are at high risk for ASCVD; their level of risk for recurrent events makes it very important that they receive appropriate optimal treatment. As high/very high risk patient population, it should be unacceptable to exclude them from the optimal care or optimal management measure: evidence-based non-statin therapies are available as well as guidance in the ACC Decision Pathway on non-statin therapies for LDL-C lowering. (Lloyd-Jones JACC, 2016: doi:10.1016/j.jacc.2016.03.519) This could impact up to 40-60% of the patients.

For many, the combination of a statin with another drug (ezetimibe) is needed to achieve the desired reduction. (Karalis Cholesterol 2012:doi.101155/2012/861924.) This is supported by the positive results of the IMProved Reduction of Outcomes: Vytorin Efficacy International Trial (IMPROVE-IT) which reported ~ 2% reduction in CV events with Vytorin (ezetimibe/simvastatin) vs. simvastatin. (Blazing Am Ht J 2014: 205-212). These results are consistent with the established relationship between absolute reductions in LDL-C and reductions in major CV events (Baigent Lancet 2010: 167-81). Murphy et al demonstrated that treatment with ezetimibe plus simvastatin resulted in a 9% reduction in total events of which 56% were first events and 44% were subsequent events. Reduction in total events was driven by reductions in MI and stroke. (Murphy JACC 2016: 353-61.) Sabatine et al reported the rate of cardiovascular events at 1 year was reduced from 2.18% in the standard-therapy group to 0.95% in the evolocumab group. (Sabatine NEJM 2015: 61-70.) Robinson et al reported a reduction in mortality as well as nonfatal events from 3.3% in the placebo group to 1.7% in the intervention group. (Robinson NEJM 2015: 1489-99.) Both studies with PCSK9 to lower LDL cholesterol levels showed approximately 50% reductions in composite CV events at 12 – 18 months.

Finally, this composite measure has been endorsed as an unweighted measure. This is the simplest approach but an all-or-none measure is usually process measures (Nolan JAMA 2006: 1168-70; Shwartz Milbank Q 2015:788-825) ; this is a combination of process and outcome/intermediate outcomes. As noted earlier, the design of the all-or-none measure implicitly assumes that each risk factor reflected in the component measures has an equal impact on the desired outcome, reduction of ASCVD risk and events. In this composite, the other component measures give very specific information that can be used to assess their impact on the ASCVD risk. However, the impact of the lipid component on risk reduction is unknown because the component measure tells us little about the statin use or the cholesterol status of the patient – we don't know what the statin dose/risk relationship is nor do we know the LDL-C level. It would not be possible to make evidence-based assumptions about the impact of this component on the risk reduction. Based on the numbers reported for this measure, it would not be possible to understand the impact of this composite measure on the success or failure of clinicians and public health efforts to reduce ASCVD events. To be useful in understanding its impact, we would recommend 2 changes.

#7 - The 2 changes needed are a) the cholesterol component needs to provide information that can be translated into an impact on ASCVD risk and b) consider modifying the measure to be a weighted measure. As Nolan points out, unweighted measures may be appropriate to get things started, before much is known about the relationship of the component parts. Not only are we past the initial stage of using the measure, given the high priority need to reduce ASCVD risk, it is a public health imperative that we move forward to ensure improving care. We have a number of ASCVD risk calculators and other research which use weighted information to assess the level of ASCVD Risk and impact of different risk factors on outcome. This information could be used to assign a weight for each of the component measures, more in line with a 'bundled measure' which is a grouping of "best practices with respect to a disease process

that individually improve care, but when applied together result in substantially greater improvement." (Shwartz Milbank Q 2015:788-825)

In summary, as written, the use of statin prescribing as the component measure to address optimal care for cholesterol management is not consistent with the evidence of 'optimal care' nor is has it been proven as written as a process measure to reduce ASCVD risk/events in RCTs or, more importantly, in the natural settings. We do not believe a single process measure is sufficient to reflect 'optimal care' or 'optimal management' for a high/very high risk person with ASCVD, as the component of the all-or-none measure addressing cholesterol management. "Optimal care" would be best demonstrated as an intermediate outcome, similar to the blood pressure and smoking component measures; then the 3 components would be closer in impact on the ASCVD risk, justifying it as an unweighted composite. In addition, instead of excluding those who are statin intolerant or do not achieve desired response with statins, it should include all patients and accept use of the evidence-based non-statins. Only when the cholesterol management component meets these criteria will it be 'optimal care' on par with the other components of the measure.

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0076

Measure Title: Optimal Vascular Care

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: 3/24/2016

Please Note:

Each of the 4 components has an individual evidence form; combined into one document. The start of each new component is labeled and highlighted in blue like: Component # 1 Blood Pressure. Updated or more recent evidence since last maintenance review in 2010/2011 is indicated in red font.

Instructions

- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- Respond to all questions as instructed with answers immediately following the question. All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 10 pages (*incudes questions/instructions*; minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- Health outcome: [3] a rationale supports the relationship of the health outcome to processes or structures of care. Applies to patient-reported outcomes (PRO), including health-related quality of life/functional status, symptom/symptom burden, experience with care, health-related behavior.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- Efficiency: [6] evidence not required for the resource use component.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the U.S. Preventive Services Task Force (USPSTF) grading definitions and methods, or Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.

5. 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 multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

Component # 1 Blood Pressure

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)
Outcome

☐ Health outcome: Click here to name the health outcome

☐Patient-reported outcome (PRO): Click here to name the PRO

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors

☒ Intermediate clinical outcome (*e.g., lab value*): Blood pressure at or below target; less than 140 systolic AND less than 90 diastolic

☐ Process: Click here to name the process

☐ Structure: Click here to name the structure

☐ Other: Click here to name what is being measured

HEALTH OUTCOME/PRO PERFORMANCE MEASURE *If not a health outcome or PRO, skip to 1a.3*

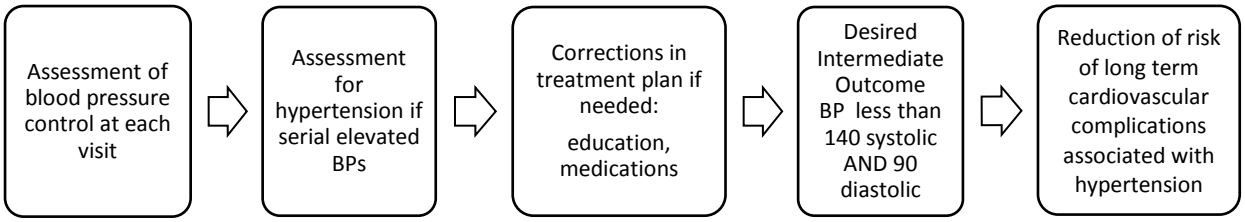
1a.2. Briefly state or diagram the path between the health outcome (or PRO) and the healthcare structures, processes, interventions, or services that influence it.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process, intervention, or service (*i.e., influence on outcome/PRO*).

Note: For health outcome/PRO performance measures, no further information is required; however, you may provide evidence for any of the structures, processes, interventions, or service identified above.

INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

1a.3. Briefly state or diagram the path between structure, process, intermediate outcome, and health outcomes. Include all the steps between the measure focus and the health outcome.



1a.3.1. What is the source of the systematic review of the body of evidence that supports the performance measure?
☒ Clinical Practice Guideline recommendation – ***complete sections 1a.4, and 1a.7***
☐ US Preventive Services Task Force Recommendation – ***complete sections 1a.5 and 1a.7***
☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*) – ***complete sections 1a.6 and 1a.7***
☐ Other – ***complete section 1a.8***

Please complete the sections indicated above for the source of evidence. You may skip the sections that do not apply.

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION

1a.4.1. Guideline citation (*including date*) and **URL for guideline** (*if available online*):

ICSI Stable Coronary Artery Disease April 2011
http://www.icsi.org/guidelines_and_more/gl_os_prot/cardiovascular/coronary_artery_disease/coronary_artery_disease__stable__3.html

ICSI Hypertension Diagnosis and Treatment November 2010
http://www.icsi.org/guidelines_and_more/gl_os_prot/cardiovascular/hypertension_4/hypertension_diagnosis_and_treatment__11.html

Treatment of Hypertension in Patients with Coronary Artery Disease
American Heart Association, American College of Cardiology, and American Society of Hypertension
Scientific Statement March 31 2015. Clive Rosendorff, Daniel T. Lackland, Matthew Allison, et. al. on behalf of the American Heart Association, American College of Cardiology, and American Society of Hypertension Hypertension. 2015;65:1372-1407. Table # 3 Summary of BP Goals (Table 3 page 1376) and Recommendations (3.3 page 1386)
http://hyper.ahajournals.org/content/65/6/1372.full.pdf+html

1a.4.2. Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

ICSI Stable Coronary Artery Disease April 2011

Address Modifiable Risk Factors and Comorbid Conditions:

Comorbid conditions that could affect myocardial ischemia may include hypertension, anemia, thyroid disease, hypoxemia and others. Modifiable risk factors for coronary heart disease need to be evaluated and may include smoking, inadequate physical activity, stress, hyperlipidemia, obesity, hypertension and diabetes mellitus.

Intervention involving any risk factor pertinent to the patient is encouraged and may include education, goal setting, and follow-up as necessary (Rutherford, 1992 [R]; Shub, 1990 [R]). Stable Coronary Artery Disease Guidelines refer to ICSI Hypertension Diagnosis and Treatment Guidelines 2010 for blood pressure goals for this population.

ICSI Hypertension Diagnosis and Treatment guideline (November 2010) for recommendations regarding blood pressure management. The recommended target blood pressure is 140/90 mmHg or less. Based on current evidence, pursuing blood pressure goals lower than < 140/90 should be considered on an individual patient basis based on clinical judgment and patient preference (ACCORD Study Group, 2010 [A], Cooper-DeHoff, 2010 [M]).

A reappraisal of evidence from randomized trials in patients with chronic heart disease or previous stroke does not show consistent evidence that cardiovascular disease risk is further reduced by more intensive lowering of blood pressure (Zanchetti, 2009 [R]). This evidence is not definitive, i.e., limitations include few trials designed to evaluate specific blood pressure goals, small differences in achieved blood pressure in many trials, and the use of active agents and corresponding placebo on top of multiple antihypertensive and other cardiovascular therapies. American Heart Association/American College of Cardiology guidelines published in 2007 called for goal office blood pressures less than 130/80 mmHg in patients with coronary disease, carotid disease, peripheral artery disease, abdominal aortic aneurysm, or a 10-year Framingham risk score of > 10% (Rosendorff, 2007 [R]). These recommendations are based on expert opinion and limited clinical evidence. A subgroup analysis of 6,400 participants of the International Verapamil SR-Trandolapril Study (INVEST) who had diabetes and coronary artery disease assessed the relationship between the degree of blood pressure control and adverse cardiovascular outcomes (Cooper-DeHoff, 2010 [M]). Tight control defined as systolic blood pressure to < 130 mmHg was not associated with fewer adverse cardiovascular outcomes compared to usual control (< 140-130 mmHg). Based on current evidence, pursuing blood pressure goals lower than < 140/90 should be considered on an individual patient basis based on clinical judgment and patient preference.

Treatment of Hypertension in Patients with Coronary Artery Disease
American Heart Association, American College of Cardiology, and American Society of Hypertension
Scientific Statement March 31 2015

Table # 3 Summary of BP Goals (Table 3 page 1376)

Please note that specifically for CAD the evidence is graded I/A.

Table 3. Summary of BP Goals

BP Goal, mm Hg	Condition	Class/Level of Evidence
<150/90	Age >80 y	IIa/B
<140/90	CAD	I/A
	ACS	IIa/C
	HF	IIa/B
<130/80	CAD	IIb/C
	Post–myocardial infarction, stroke or TIA, carotid artery disease, PAD, AAA	IIb/C

AAA indicates abdominal aortic aneurysm; ACS, acute coronary syndrome; BP, blood pressure; CAD, coronary artery disease; HF, heart failure; PAD, peripheral arterial disease; and TIA, transient ischemic attack.

Additional Recommendations for Blood Pressure Targets: (Table 3.3 pg. 1386)

1. The <140/90-mm Hg BP target is reasonable for the secondary prevention of cardiovascular events in patients with hypertension and CAD (Class IIa; Level of Evidence B).
2. A lower target BP (<130/80 mm Hg) may be appropriate in some individuals with CAD, previous MI, stroke or transient ischemic attack, or CAD risk equivalents (carotid artery disease, PAD, abdominal aortic aneurysm) (Class IIb; Level of Evidence B)

1a.4.3. Grade assigned to the quoted recommendation <u>with definition</u> of the grade:

Grade IA
Class I = Benefit >>> outweighs risk. Procedure/ Treatment SHOULD be performed/ administered
Level A = Multiple populations evaluated. Data derived from multiple randomized clinical trials or meta-analyses. Recommendation that procedure or treatment is useful/ effective. Sufficient evidence from multiple randomized control trials or meta-analysis.

1a.4.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: If separate grades for the strength of the evidence, report them in section 1a.7.*)

Table 1. Applying Classification of Recommendations and Levels of Evidence.

ESTIMATE OF CERTAINTY (PRECISION) OF TREATMENT EFFECT	SIZE OF TREATMENT EFFECT: CLASS I *Benefit >>> Risk* Procedure/Treatment **SHOULD** be performed/ administered	CLASS IIa *Benefit >> Risk* *Additional studies with focused objectives needed* **IT IS REASONABLE** to perform procedure/administer treatment	CLASS IIb *Benefit ≥ Risk* *Additional studies with broad objectives needed; additional registry data would be helpful* Procedure/Treatment **MAY BE CONSIDERED**	CLASS III *No Benefit* or CLASS III *Harm* COR III: No benefit — Procedure/Test: Not Helpful; Treatment: No Proven Benefit COR III: Harm — Procedure/Test: Excess Cost w/o Benefit or Harmful; Treatment: Harmful to Patients
LEVEL A Multiple populations evaluated* Data derived from multiple randomized clinical trials or meta-analyses	■ Recommendation that procedure or treatment is useful/effective ■ Sufficient evidence from multiple randomized trials or meta-analyses	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Sufficient evidence from multiple randomized trials or meta-analyses
LEVEL B Limited populations evaluated* Data derived from a single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is useful/effective ■ Evidence from single randomized trial or nonrandomized studies	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Evidence from single randomized trial or nonrandomized studies
LEVEL C Very limited populations evaluated* Only consensus opinion of experts, case studies, or standard of care	■ Recommendation that procedure or treatment is useful/effective ■ Only expert opinion, case studies, or standard of care	■ Recommendation in favor of treatment or procedure being useful/effective ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation's usefulness/efficacy less well established ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Only expert opinion, case studies, or standard of care
Suggested phrases for writing recommendations	should is recommended is indicated is useful/effective/beneficial	is reasonable can be useful/effective/beneficial is probably recommended or indicated	may/might be considered may/might be reasonable usefulness/effectiveness is unknown/unclear/uncertain or not well established	COR III: No Benefit: is not recommended; is not indicated; should not be performed/ administered/ other; is not useful/ beneficial/ effective COR III: Harm: potentially harmful; causes harm; associated with excess morbidity/mortality; should not be performed/ administered/ other
Comparative effectiveness phrases†	treatment/strategy A is recommended/indicated in preference to treatment B treatment A should be chosen over treatment B	treatment/strategy A is probably recommended/indicated in preference to treatment B it is reasonable to choose treatment A over treatment B		

A recommendation with Level of Evidence B or C does not imply that the recommendation is weak. Many important clinical questions addressed in the guidelines do not lend themselves to clinical trials. Although randomized trials are unavailable, there may be a very clear clinical consensus that a particular test or therapy is useful or effective.

*Data available from clinical trials or registries about the usefulness/efficacy in different subpopulations, such as sex, age, history of diabetes, history of prior myocardial infarction, history of heart failure, and prior aspirin use.

†For comparative effectiveness recommendations (Class I and IIa; Level of Evidence A and B only), studies that support the use of comparator verbs should involve direct comparisons of the treatments or strategies being evaluated.

1a.4.5. Citation and URL for methodology for grading recommendations (*if different from 1a.4.1*)**:**

1a.4.6. If guideline is evidence-based (rather than expert opinion), are the details of the quantity, quality, and consistency of the body of evidence available (e.g., evidence tables)?

☒ Yes → ***complete section 1a.7***

☐ No → ***report on another systematic review of the evidence in sections 1a.6 and 1a.7; if another review does not exist, provide what is known from the guideline review of evidence in 1a.7***

1a.5. UNITED STATES PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

1a.5.1. Recommendation citation (*including date*) and **URL for recommendation** (*if available online*):

1a.5.2. Identify recommendation number and/or page number and **quote verbatim, the specific recommendation**.

1a.5.3. Grade assigned to the quoted recommendation with definition of the grade:

1a.5.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: the grading system for the evidence should be reported in section 1a.7.*)

1a.5.5. Citation and URL for methodology for grading recommendations (*if different from 1a.5.1*)**:**

Complete section 1a.7

1a.6. OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1. Citation (*including date*) and **URL** (*if available online*):

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

Complete section 1a.7

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

If more than one systematic review of the evidence is identified above, you may choose to summarize the one (or more) for which the best information is available to provide a summary of the quantity, quality, and consistency of the body of evidence. Be sure to identify which review is the basis of the responses in this section and if more than one, provide a separate response for each review.

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

Treatment of Hypertension in Patients With Coronary Artery Disease
American Heart Association, American College of Cardiology, and American Society of Hypertension
Scientific Statement March 31 2015.

The treatment of hypertension for patients with cardiovascular disease to a target BP goal of less than 140 systolic and less than 90 diastolic.

1a.7.2. Grade assigned for the quality of the quoted evidence with definition of the grade:

Grade IA

Class I = Benefit >>> outweighs risk. Procedure/ Treatment SHOULD be performed/ administered

Level A = Multiple populations evaluated. Data derived from multiple randomized clinical trials or meta-analyses. Recommendation that procedure or treatment is useful/ effective. Sufficient evidence from multiple randomized control trials or meta-analysis.

1a.7.3. Provide all other grades and associated definitions for strength of the evidence in the grading system.

Table 1. Applying Classification of Recommendations and Levels of Evidence.

SIZE OF TREATMENT EFFECT

ESTIMATE OF CERTAINTY (PRECISION) OF TREATMENT EFFECT	CLASS I *Benefit >>> Risk* Procedure/Treatment **SHOULD** be performed/ administered	CLASS IIa *Benefit >> Risk* *Additional studies with focused objectives needed* **IT IS REASONABLE** to perform procedure/administer treatment	CLASS IIb *Benefit ≥ Risk* *Additional studies with broad objectives needed; additional registry data would be helpful* Procedure/Treatment **MAY BE CONSIDERED**	CLASS III *No Benefit* or CLASS III *Harm* Procedure/ Test / Treatment COR III: No benefit — Not Helpful — No Proven Benefit COR III: Harm — Excess Cost w/o Benefit or Harmful — Harmful to Patients
LEVEL A Multiple populations evaluated* Data derived from multiple randomized clinical trials or meta-analyses	■ Recommendation that procedure or treatment is useful/effective ■ Sufficient evidence from multiple randomized trials or meta-analyses	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Sufficient evidence from multiple randomized trials or meta-analyses
LEVEL B Limited populations evaluated* Data derived from a single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is useful/effective ■ Evidence from single randomized trial or nonrandomized studies	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Evidence from single randomized trial or nonrandomized studies
LEVEL C Very limited populations evaluated* Only consensus opinion of experts, case studies, or standard of care	■ Recommendation that procedure or treatment is useful/effective ■ Only expert opinion, case studies, or standard of care	■ Recommendation in favor of treatment or procedure being useful/effective ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation's usefulness/efficacy less well established ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Only expert opinion, case studies, or standard of care
Suggested phrases for writing recommendations	should is recommended is indicated is useful/effective/beneficial	is reasonable can be useful/effective/beneficial is probably recommended or indicated	may/might be considered may/might be reasonable usefulness/effectiveness is unknown/unclear/uncertain or not well established	COR III: No Benefit is not recommended is not indicated should not be performed/ administered/ other is not useful/ beneficial/ effective COR III: Harm potentially harmful causes harm associated with excess morbidity/mortality should not be performed/ administered/ other
Comparative effectiveness phrases†	treatment/strategy A is recommended/indicated in preference to treatment B treatment A should be chosen over treatment B	treatment/strategy A is probably recommended/indicated in preference to treatment B it is reasonable to choose treatment A over treatment B		

A recommendation with Level of Evidence B or C does not imply that the recommendation is weak. Many important clinical questions addressed in the guidelines do not lend themselves to clinical trials. Although randomized trials are unavailable, there may be a very clear clinical consensus that a particular test or therapy is useful or effective.

*Data available from clinical trials or registries about the usefulness/efficacy in different subpopulations, such as sex, age, history of diabetes, history of prior myocardial infarction, history of heart failure, and prior aspirin use.

†For comparative effectiveness recommendations (Class I and IIa; Level of Evidence A and B only), studies that support the use of comparator verbs should involve direct comparisons of the treatments or strategies being evaluated.

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range: 1996 to 2015

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

From the section on determining BP Goals pg. 1382 to 1386

8 randomized control trials, 6 prospective observational studies, 1 meta-analysis including 147 RCTs and 1 meta-regression including 31 interventional trials.

1a.7.6. What is the overall quality of evidence across studies in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

Conclusions from: Treatment of Hypertension in Patients with Coronary Artery Disease
American Heart Association, American College of Cardiology, and American Society of Hypertension
Scientific Statement March 31 2015.

Lower SBP values may be associated with better stroke outcomes except in the case of PROFESS, and the evidence for CAD outcomes is equivocal. The evidence that excessive lowering of DBP may compromise cardiac outcomes (the J curve) is inconsistent. Epidemiological and clinical trial evidence both support and refute the existence of a J curve for DBP but not SBP, which suggests the presence of major confounders of data interpretation, including selection bias, comorbidities, and nonlinear interactions among age, decreasing DBP, and increasing cardiovascular risk. The vast majority of hypertensive individuals, including those with overt cardiac disease, will not experience problems related to lowering of DBP when standard antihypertensive medications are used. Concerns that coronary perfusion is limited by an autoregulatory threshold have not yet been validated in humans with healthy or even diseased coronary arteries, and no consensus exists on the minimum safe level of DBP in these

individuals. Although an autoregulatory threshold has not been defined in humans, with or without CAD, it is clear, mainly from ACCORD, that lower BP targets, down to levels <120/80 mm Hg, protect against stroke and do not significantly increase CAD events. Most studies that have addressed lower BP targets have achieved DBP values in the 70- to 79-mm Hg range, which appears to be safe.

Therefore, a reasonable recommendation would be a BP target of <140/90 mm Hg for the secondary prevention of cardiovascular events in patients with CAD. However, there are some epidemiological data, several post hoc analyses of clinical trials, and a plethora of other data that support, but do not prove, that a lower target (<130/80 mm Hg) may be appropriate in some individuals with CAD. We counsel that the BP should be lowered slowly in patients with occlusive CAD with evidence of myocardial ischemia, and caution is advised in inducing decreases in DBP to <60 mm Hg, particularly if the patient is >60 years of age. In older hypertensive individuals with wide pulse pressures, lowering SBP may cause very low DBP values (<60 mm Hg). This should alert the clinician to assess carefully any untoward signs or symptoms, especially those resulting from myocardial ischemia. In patients >80 years of age, a reasonable BP target is <150/80 mm Hg, although there are no direct data to support this, or any other specific BP goal, in this age group.

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) across studies in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

Please see conclusions from the guideline group in section 1a.7.6.

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

Please see conclusions from the guideline group in section 1a.7.6.

The estimate of benefit outweighs significantly with a Class I Level A grade recommendation.

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for each new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

1a.8 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.8.1 What process was used to identify the evidence?

1a.8.2. Provide the citation and summary for each piece of evidence.

Component # 2 Stain Medication

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)
Outcome

☒ Health outcome: Click here to name the health outcome

☐Patient-reported outcome (PRO): Click here to name the PRO

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors

☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☒ Process: Appropriate use of statin medication (if patient with ischemic vascular disease is without contraindications or exceptions to statin use)

☐ Structure: Click here to name the structure

☐ Other: Click here to name what is being measured

HEALTH OUTCOME/PRO PERFORMANCE MEASURE *If not a health outcome or PRO, skip to 1a.3*

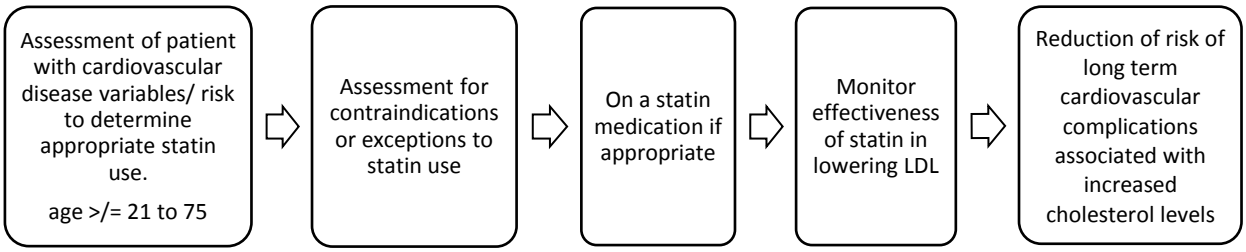
1a.2. Briefly state or diagram the path between the health outcome (or PRO) and the healthcare structures, processes, interventions, or services that influence it.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process, intervention, or service (*i.e., influence on outcome/PRO*).

Note: For health outcome/PRO performance measures, no further information is required; however, you may provide evidence for any of the structures, processes, interventions, or service identified above.

INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

1a.3. Briefly state or diagram the path between structure, process, intermediate outcome, and health outcomes. Include all the steps between the measure focus and the health outcome.



1a.3.1. What is the source of the systematic review of the body of evidence that supports the performance measure?

☒ Clinical Practice Guideline recommendation – ***complete sections 1a.4, and 1a.7***

☐ US Preventive Services Task Force Recommendation – ***complete sections 1a.5 and 1a.7***

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*) – ***complete sections 1a.6 and 1a.7***

☐ Other – ***complete section 1a.8***

Please complete the sections indicated above for the source of evidence. You may skip the sections that do not apply.

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION

1a.4.1. Guideline citation (*including date*) and **URL for guideline** (*if available online*):

ICSI Stable Coronary Artery Disease April 2011
http://www.icsi.org/guidelines_and_more/gl_os_prot/cardiovascular/coronary_artery_disease/coronary_artery_disease__stable__3.html

ICSI Lipid Management in Adults October 2009
http://www.icsi.org/guidelines_and_more/gl_os_prot/cardiovascular/lipid_management_3/lipid_management_in_adults__4.html

ICSI Lipid Management in Adults November 2013
http://www.icsi.org/guidelines__more/catalog_guidelines_and_more/catalog_guidelines/catalog_cardiovascular_guidelines/lipid/
2013 ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adult
https://circ.ahajournals.org/content/early/2013/11/11/01.cir.0000437738.63853.7a.full.pdf+html

1a.4.2. Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

ICSI Stable Coronary Artery Disease April 2011
Address Modifiable Risk Factors and Comorbid Conditions:
Comorbid conditions that could affect myocardial ischemia may include hypertension, anemia, thyroid disease, hypoxemia and others. Modifiable risk factors for coronary heart disease need to be evaluated and may include smoking, inadequate physical activity, stress, hyperlipidemia, obesity, hypertension and diabetes mellitus.
Intervention involving any risk factor pertinent to the patient is encouraged and may include education, goal setting, and follow-up as necessary (Rutherford, 1992 [R]; Shub, 1990 [R]).

Hyperlipidemia:
A fasting lipid profile should be evaluated for appropriate patients with stable coronary artery disease. Secondary prevention is important in these patients, who should be treated aggressively for hyperlipidemia. Many patients will require both pharmacologic and non-pharmacologic interventions to reach target goals.
Target goals for hyperlipidemic patients with coronary artery disease include:
LDL – less than 100 mg/dL
HDL – 40 mg/dL or greater
Triglycerides – less than 150 mg/dL
(ALLHAT, 2002 [A]; Cannon, 2004 [A]; Downs, 1998 [A]; Heart Protection Study Collaborative Group, 2002 [A]; LaRosa, 1999 [M]; Lipid Research Clinics Program, 1984 [A]; Nissen, 2004 [A]; Pignone, 2000 [M]; Sever, 2003 [A]; Shepherd, 2002 [A]; Shepherd,1995 [A]; Topol, 2004 [R]; Goldberg, 1998 [A]; LIPID Study Group, 1998 [A]; Scandinavian Simvastatin Survival Study Group, 1994) [A].
Please also refer to the ICSI Lipid Management in Adults Guideline

ICSI Lipid Management in Adults (updated Nov 2013/ completed prior to ACC/AHA release)

Initiate Statin Treatment Recommendations:

Clinicians should initiate statin therapy regardless of LDL, in patients with established ASCVD (Strong Recommendation, High Quality Evidence) (Cannon, 2004; Heart Protection Study Collaborative Group, 2002; Shepard, 2002; La Rosa, 1999; LIPID Study Group, 1998; Goldberg, 1998; Scandinavian Simvastatin Survival Study Group, 1994).

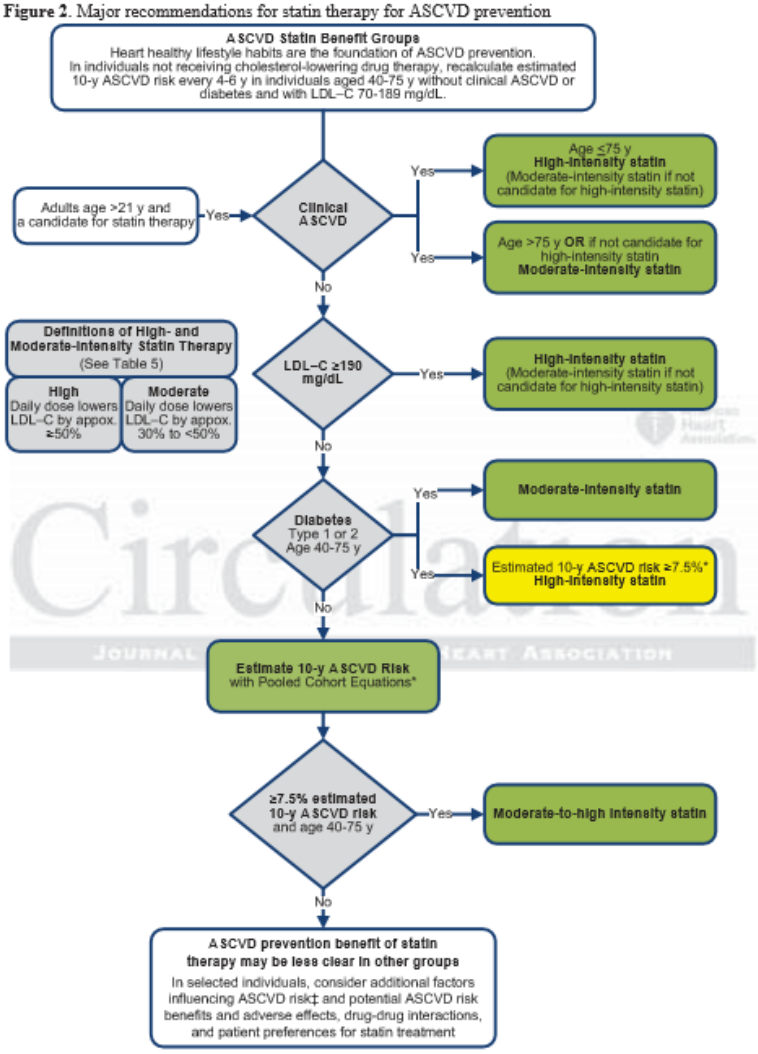
2013 ACC/AHA Guideline: Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults

2.2. Four Major Statin Benefit Groups (pg 13- 15)
The Expert Panel found extensive and consistent evidence supporting the use of statins for the prevention of ASCVD in many higher risk primary and all secondary prevention individuals without NYHA class II-IV heart failure and who were not receiving hemodialysis. In the RCTs reviewed, initiation of moderate intensity therapy (lowering LDL–C by approximately 30% to <50%), or high-intensity statin therapy (lowering LDL–C by approximately •50%), is a critical factor in reducing ASCVD events. Moreover, statin therapy reduces ASCVD events across the spectrum of baseline LDL–C levels >70 mg/dL. In addition, the relative reduction in ASCVD risk is consistent for primary and secondary prevention and for various patient subgroups. Of note, the absolute reduction in ASCVD events is proportional to baseline absolute ASCVD risk. Therefore, statin therapy is recommended for individuals at increased ASCVD risk who are most likely to experience a net benefit in terms of the potential for ASCVD risk reduction and the potential for adverse effects. On the basis of this large and consistent body of evidence, 4 major statin benefit groups were identified for whom the ASCVD risk reduction clearly outweighs the risk of adverse events. Individuals 1) with clinical ASCVD, 2) primary elevations of LDL–C >190 mg/dL, 3) diabetes aged 40 to 75 years with LDL– C 70 to189 mg/dL and without clinical ASCVD, or 4) without clinical ASCVD or diabetes with LDL–C 70 to189 mg/dL and estimated 10-year ASCVD risk >7.5%. Clinical ASCVD is defined by the inclusion criteria for

the secondary prevention statin RCTs (acute coronary syndromes, or a history of MI, stable or unstable angina, coronary or other arterial revascularization, stroke, TIA, or peripheral arterial disease presumed to be of atherosclerotic origin).

Stone NJ, et al.
2013 ACC/AHA Blood Cholesterol Guideline

Figure 2. Major recommendations for statin therapy for ASCVD prevention



Page 15

Statin Treatment Recommendations (Section 4)

Recommendations for Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults—Statin Treatment (Table 4 page 22)

- High-intensity statin therapy should be initiated or continued as first-line therapy in women and men less than or equal to age 75 who have clinical ASCVD*, unless contraindicated.
- In individuals with clinical ASCVD* in whom high-intensity statin therapy would otherwise be used, when high-intensity statin therapy is contraindicated or when characteristics predisposing to statin-associated adverse effects are present, moderate-intensity statin should be used as the second option if tolerated (Table 8 for Safety of Statins, Recommendation 1).

*Clinical ASCVD includes acute coronary syndromes, history of MI, stable or unstable angina, coronary or other arterial revascularization, stroke, TIA, or peripheral arterial disease presumed to be of atherosclerotic origin.

Table 4. Recommendations for Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults—Statin Treatment
(High-, moderate-, and low-statin intensities are defined in Table 5)

Recommendations	NHLBI Grade	NHLBI Evidence Statements	ACC/AHA COR	ACC/AHA LOE
Secondary Prevention				
1. High-intensity statin therapy should be initiated or continued as first-line therapy in women and men ≤75 years of age who have *clinical ASCVD**, unless contraindicated.	A (Strong)	1, 6-8, 10-23, 26-28	I	A
2. In individuals with *clinical ASCVD** in whom high-intensity statin therapy would otherwise be used, when high-intensity statin therapy is contraindicated† or when characteristics predisposing to statin-associated adverse effects are present, moderate-intensity statin should be used as the second option if tolerated (Table 8 for Safety of Statins, Recommendation 1).	A (Strong)	13-22, 24, 27, 28	I	A

4.3. Secondary Prevention (page 26-27)

Women and men with clinical ASCVD (defined from the RCT inclusion criteria as acute coronary syndromes; history of MI, stable or unstable angina, coronary revascularization, stroke, or TIA presumed to be of atherosclerotic origin, and peripheral arterial disease or revascularization) are at increased risk for recurrent ASCVD and ASCVD death. An extensive body of evidence demonstrates that high-intensity statin therapy reduces ASCVD events more than moderate-intensity statin therapy (Table 4) in individuals with clinical ASCVD. High-intensity statin therapy should be initiated for adults less than or equal to 75 years of age with clinical ASCVD who are not receiving statin therapy or the intensity should be increased in those receiving a low- or moderate-intensity statin, unless they have a history of intolerance to high-intensity statin therapy or other characteristics that may influence safety (Section 5). This is consistent with RCT data. In 2 trials (46,47), patients were previously treated with a moderately intensive statin and in 2 trials 75% to 97% of patients had not received prior statin therapy (48,79). The high-intensity statins atorvastatin 80 mg and rosuvastatin 20 mg daily reduce LDL–C greater than 50% on average and have been shown to reduce ASCVD events in RCTs. Although atorvastatin 40 mg reduces LDL by approximately greater than or equal to 50%, this dose was only used in 1 RCT if the participant was unable to tolerate atorvastatin 80 mg/dL. Whether an individual receiving atorvastatin 40 mg should be uptitrated to atorvastatin 80 mg should be based the potential for an ASCVD risk reduction benefit and the potential for adverse effects (including drug-drug interactions), as well as patient preferences. In individuals with clinical ASCVD in whom high-intensity statin therapy would otherwise be used, when either high-intensity statin therapy is contraindicated or when characteristics predisposing to statin associated adverse effects are present, moderate-intensity statin should be used as the second option, if tolerated (Section 5). In the relatively few individuals >75 years of age who were included in RCTs of high- versus moderate-intensity statin therapy there was not clear evidence of an additional reduction in ASCVD events from high-intensity statin therapy.

1a.4.3. Grade assigned to the quoted recommendation <u>with definition</u> of the grade:

Grade IA

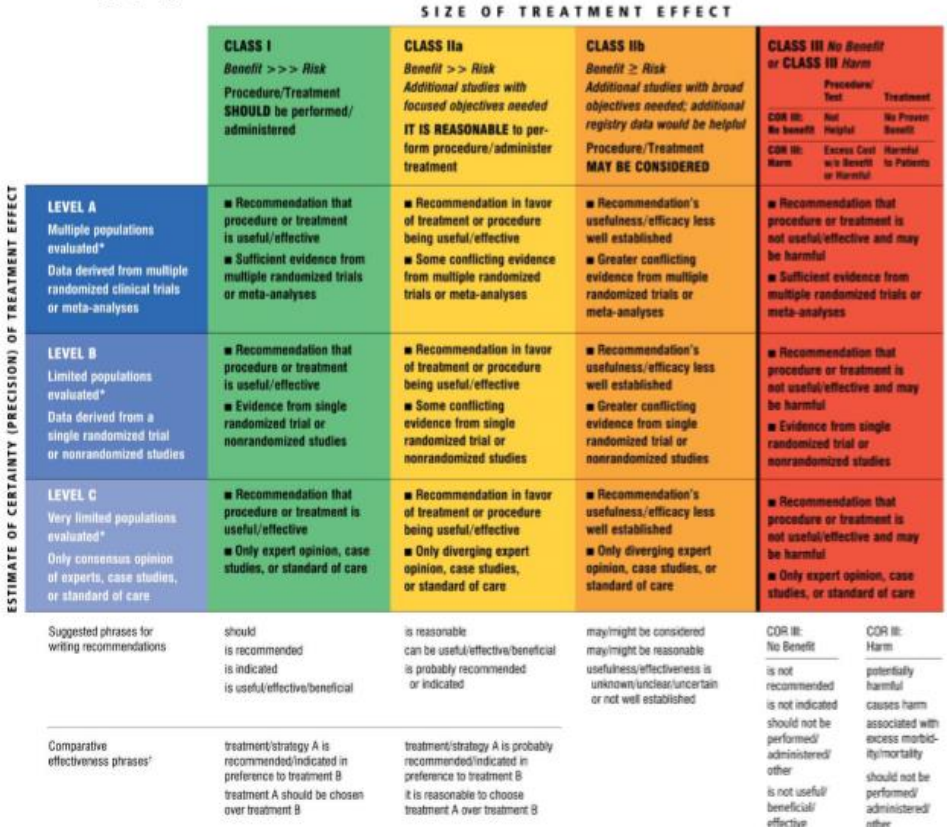
Class I = Benefit >>> outweighs risk. Procedure/ Treatment SHOULD be performed/ administered

Level A = Multiple populations evaluated. Data derived from multiple randomized clinical trials or meta-analyses. Recommendation that procedure or treatment is useful/ effective. Sufficient evidence from multiple randomized control trials or meta-analysis.

1a.4.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: If separate grades for the strength of the evidence, report them in section 1a.7.*)

ACC/AHA guidelines. Each recommendation has been mapped from the NHLBI grading format to the ACC/AHA Class of Recommendation/Level of Evidence (COR/LOE) construct (Table 1) and is expressed in both formats. Because of the inherent differences in grading systems and the clinical questions driving the recommendations, alignment between the NHLBI and ACC/AHA formats is in some cases imperfect. Explanations of these variations are noted in the recommendation tables, where applicable.

Table 1. Applying Classification of Recommendation and Level of Evidence



ESTIMATE OF CERTAINTY (PRECISION) OF TREATMENT EFFECT	SIZE OF TREATMENT EFFECT: CLASS I *Benefit >>> Risk* Procedure/Treatment **SHOULD** be performed/ administered	CLASS IIa *Benefit >> Risk* *Additional studies with focused objectives needed* **IT IS REASONABLE** to perform procedure/administer treatment	CLASS IIb *Benefit ≥ Risk* *Additional studies with broad objectives needed; additional registry data would be helpful* Procedure/Treatment **MAY BE CONSIDERED**	CLASS III *No Benefit* or CLASS III *Harm* Procedure/ Test / Treatment COR III: No benefit — Not Helpful — No Proven Benefit COR III: Harm — Excess Cost w/o Benefit or Harmful — Harmful to Patients
LEVEL A Multiple populations evaluated* Data derived from multiple randomized clinical trials or meta-analyses	■ Recommendation that procedure or treatment is useful/effective ■ Sufficient evidence from multiple randomized trials or meta-analyses	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Sufficient evidence from multiple randomized trials or meta-analyses
LEVEL B Limited populations evaluated* Data derived from a single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is useful/effective ■ Evidence from single randomized trial or nonrandomized studies	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Evidence from single randomized trial or nonrandomized studies
LEVEL C Very limited populations evaluated* Only consensus opinion of experts, case studies, or standard of care	■ Recommendation that procedure or treatment is useful/effective ■ Only expert opinion, case studies, or standard of care	■ Recommendation in favor of treatment or procedure being useful/effective ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation's usefulness/efficacy less well established ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Only expert opinion, case studies, or standard of care
Suggested phrases for writing recommendations	should is recommended is indicated is useful/effective/beneficial	is reasonable can be useful/effective/beneficial is probably recommended or indicated	may/might be considered may/might be reasonable usefulness/effectiveness is unknown/unclear/uncertain or not well established	COR III: No Benefit: is not recommended; is not indicated; should not be performed/ administered/ other; is not useful/ beneficial/ effective COR III: Harm: potentially harmful; causes harm; associated with excess morbidity/mortality; should not be performed/ administered/ other
Comparative effectiveness phrases†	treatment/strategy A is recommended/indicated in preference to treatment B treatment A should be chosen over treatment B	treatment/strategy A is probably recommended/indicated in preference to treatment B it is reasonable to choose treatment A over treatment B		

1a.4.5. Citation and URL for methodology for grading recommendations (*if different from 1a.4.1*)**:**

1a.4.6. If guideline is evidence-based (rather than expert opinion), are the details of the quantity, quality, and consistency of the body of evidence available (e.g., evidence tables)?

☒ Yes → ***complete section 1a.7***

☐ No → ***report on another systematic review of the evidence in sections 1a.6 and 1a.7; if another review does not exist, provide what is known from the guideline review of evidence in 1a.7***

1a.5. UNITED STATES PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

1a.5.1. Recommendation citation (*including date*) and **URL for recommendation** (*if available online*):

1a.5.2. Identify recommendation number and/or page number and **quote verbatim, the specific recommendation**.

1a.5.3. Grade assigned to the quoted recommendation with definition of the grade:

1a.5.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: the grading system for the evidence should be reported in section 1a.7.*)

1a.5.5. Citation and URL for methodology for grading recommendations (*if different from 1a.5.1*)**:**

Complete section 1a.7

_____

1a.6. OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1. Citation (*including date*) and **URL** (*if available online*):

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

Complete section 1a.7

_____

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

If more than one systematic review of the evidence is identified above, you may choose to summarize the one (or more) for which the best information is available to provide a summary of the quantity, quality, and consistency of the body of evidence. Be sure to identify which review is the basis of the responses in this section and if more than one, provide a separate response for each review.

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adult (Nov 2013)

The prevention of secondary cardiovascular events for patients with cardiovascular disease by the appropriate prescribing of statin medications.

1a.7.2. Grade assigned for the quality of the quoted evidence with definition of the grade:

Grade IA

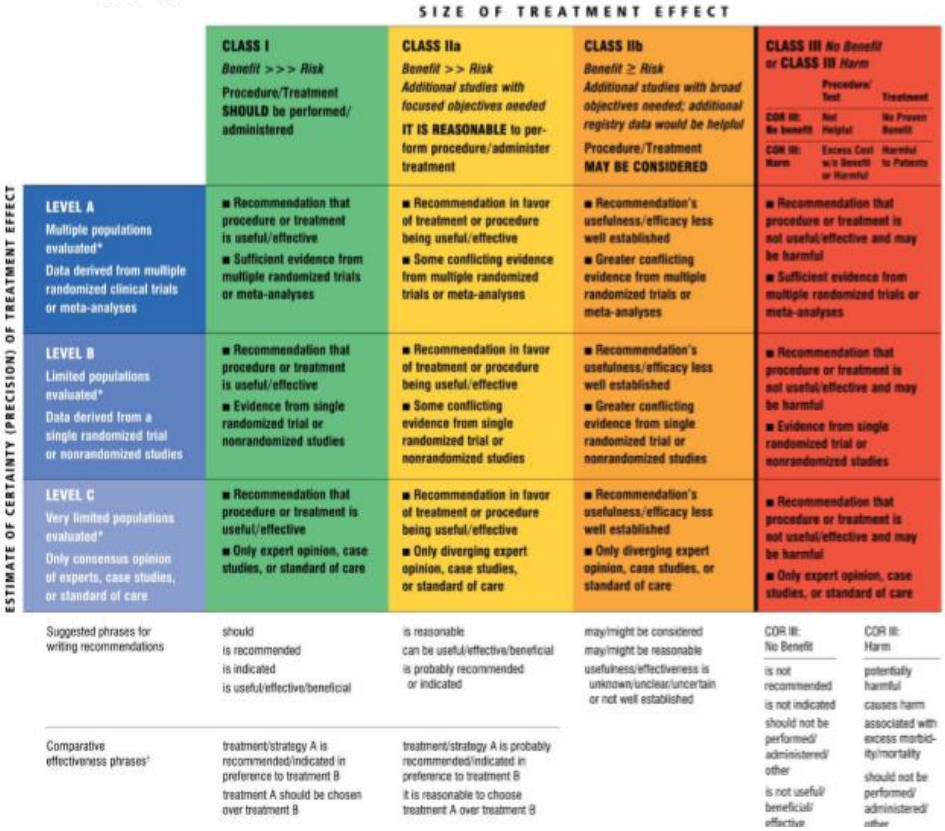
Class I = Benefit >>> outweighs risk. Procedure/ Treatment SHOULD be performed/ administered

Level A = Multiple populations evaluated. Data derived from multiple randomized clinical trials or meta-analyses. Recommendation that procedure or treatment is useful/ effective. Sufficient evidence from multiple randomized control trials or meta-analysis.

1a.7.3. Provide all other grades and associated definitions for strength of the evidence in the grading system.

ACC/AHA guidelines. Each recommendation has been mapped from the NHLBI grading format to the ACC/AHA Class of Recommendation/Level of Evidence (COR/LOE) construct (Table 1) and is expressed in both formats. Because of the inherent differences in grading systems and the clinical questions driving the recommendations, alignment between the NHLBI and ACC/AHA formats is in some cases imperfect. Explanations of these variations are noted in the recommendation tables, where applicable.

Table 1. Applying Classification of Recommendation and Level of Evidence



ESTIMATE OF CERTAINTY (PRECISION) OF TREATMENT EFFECT	SIZE OF TREATMENT EFFECT: CLASS I *Benefit >>> Risk* Procedure/Treatment **SHOULD** be performed/ administered	CLASS IIa *Benefit >> Risk* *Additional studies with focused objectives needed* **IT IS REASONABLE** to perform procedure/administer treatment	CLASS IIb *Benefit ≥ Risk* *Additional studies with broad objectives needed; additional registry data would be helpful* Procedure/Treatment **MAY BE CONSIDERED**	CLASS III *No Benefit* or CLASS III *Harm* COR III: No benefit — Procedure/Test: Not Helpful; Treatment: No Proven Benefit COR III: Harm — Procedure/Test: Excess Cost w/o Benefit or Harmful; Treatment: Harmful to Patients
LEVEL A Multiple populations evaluated* Data derived from multiple randomized clinical trials or meta-analyses	■ Recommendation that procedure or treatment is useful/effective ■ Sufficient evidence from multiple randomized trials or meta-analyses	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Sufficient evidence from multiple randomized trials or meta-analyses
LEVEL B Limited populations evaluated* Data derived from a single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is useful/effective ■ Evidence from single randomized trial or nonrandomized studies	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Evidence from single randomized trial or nonrandomized studies
LEVEL C Very limited populations evaluated* Only consensus opinion of experts, case studies, or standard of care	■ Recommendation that procedure or treatment is useful/effective ■ Only expert opinion, case studies, or standard of care	■ Recommendation in favor of treatment or procedure being useful/effective ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation's usefulness/efficacy less well established ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Only expert opinion, case studies, or standard of care
Suggested phrases for writing recommendations	should is recommended is indicated is useful/effective/beneficial	is reasonable can be useful/effective/beneficial is probably recommended or indicated	may/might be considered may/might be reasonable usefulness/effectiveness is unknown/unclear/uncertain or not well established	COR III: No Benefit is not recommended is not indicated should not be performed/ administered/ other is not useful/ beneficial/ effective COR III: Harm potentially harmful causes harm associated with excess morbidity/mortality should not be performed/ administered/ other
Comparative effectiveness phrases†	treatment/strategy A is recommended/indicated in preference to treatment B treatment A should be chosen over treatment B	treatment/strategy A is probably recommended/indicated in preference to treatment B it is reasonable to choose treatment A over treatment B		

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range: 1998 to 2013

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

60 randomized control trials, 1 systematic review and 1 meta-analysis.

1a.7.6. What is the overall quality of evidence across studies in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

One of the critical questions that the guideline workgroup included in their study related to the previous practice recommendation of treating to an LDL target.

CQ1: What is the evidence for LDL–C and non-HDL–C goals for the secondary prevention of ASCVD?

The Expert Panel reviewed 19 RCTs to answer CQ1. Although supported conceptually by an extrapolation of observational studies and observational data from RCTs, no data were identified regarding treatment or titration to a specific LDL–C goal in adults with clinical ASCVD. The majority of studies confirming the efficacy of cholesterol reduction in improving clinical outcomes in patients with clinical ASCVD used a single fixed-dose statin therapy to lower LDL–C levels. In the 4S trial, 37% had the dose of simvastatin raised from 20 mg to 40 mg per day to achieve a total cholesterol level <200 mg/day. The Expert Panel was unable to find any RCTs

that evaluated titration of all individuals in a treatment group to specific LDL–C targets <100 mg/dL or <70 mg/dL. Nor were any RCTs comparing 2 LDL–C treatment targets identified. No statin RCTs reporting on-treatment non-HDL–C levels were identified.

The quality of evidence across studies related to the secondary prevention of cardiovascular events was rated as strong. [Level A = Multiple populations evaluated. Data derived from multiple randomized clinical trials or meta-analyses. Recommendation that procedure or treatment is useful/ effective. Sufficient evidence from multiple randomized control trials or meta-analysis.]

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) across studies in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

Please see conclusions from the guideline group in section 1a.7.6.

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

Please see conclusions from the guideline group in section 1a.7.6.

The estimate of benefit outweighs significantly with a Class I Level A grade recommendation.

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for each new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

1a.8 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.8.1 What process was used to identify the evidence?

1a.8.2. Provide the citation and summary for each piece of evidence.

Component # 3 Tobacco Free

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)
Outcome

☒ Health outcome: Patient is tobacco-free

☐Patient-reported outcome (PRO): Click here to name the PRO

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors

☐ Intermediate clinical outcome (*e.g., lab value*):

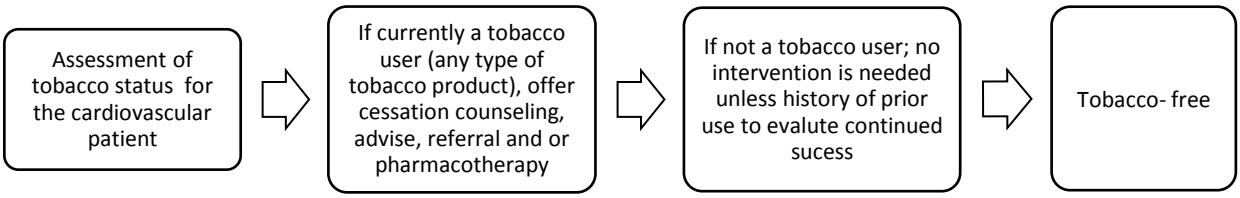
☐ Process: Click here to name the process

☐ Structure: Click here to name the structure

☐ Other: Click here to name what is being measured

HEALTH OUTCOME/PRO PERFORMANCE MEASURE *If not a health outcome or PRO, skip to 1a.3*

1a.2. Briefly state or diagram the path between the health outcome (or PRO) and the healthcare structures, processes, interventions, or services that influence it.



1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process, intervention, or service (*i.e., influence on outcome/PRO*).

ICSI Stable Coronary Artery Disease April 2011

Address Modifiable Risk Factors and Comorbid Conditions:

Comorbid conditions that could affect myocardial ischemia may include hypertension, anemia, thyroid disease, hypoxemia and others. Modifiable risk factors for coronary heart disease need to be evaluated and may include smoking, inadequate physical activity, stress, hyperlipidemia, obesity, hypertension and diabetes mellitus.

Intervention involving any risk factor pertinent to the patient is encouraged and may include education, goal setting, and follow-up as necessary (Rutherford, 1992 [R]; Shub, 1990 [R]).

According to the United States Preventive Services Task Force (USPSTF) cigarette smoking and exposure to smoke result in more than 480 000 premature deaths in the United States every year, along with substantial illness . Despite considerable progress in tobacco control over the past 50 years, in 2013, an estimated 17.8% of U.S. adults (3) and 15.9% of pregnant women aged 15 to 44 years were current cigarette smokers. The Centers for Disease Control indicates that smoking is a major cause of cardiovascular disease and that tobacco use contributes to heart disease and stroke by raising triglycerides, lowering (good) HDL cholesterol, increases clotting factors, damages cells that line blood vessels, increases the buildup of plague, and causes thickening and narrowing of blood vessels.

http://www.cdc.gov/tobacco/basic_information/health_effects/heart_disease/

Note: For health outcome/PRO performance measures, no further information is required; however, you may provide evidence for any of the structures, processes, interventions, or service identified above.

INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

1a.3. Briefly state or diagram the path between structure, process, intermediate outcome, and health outcomes. Include all the steps between the measure focus and the health outcome.

1a.3.1. What is the source of the systematic review of the body of evidence that supports the performance measure?
☐ Clinical Practice Guideline recommendation – ***complete sections 1a.4, and 1a.7***
☐ US Preventive Services Task Force Recommendation – ***complete sections 1a.5 and 1a.7***
☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*) – ***complete sections 1a.6 and 1a.7***
☐ Other – ***complete section 1a.8***

Please complete the sections indicated above for the source of evidence. You may skip the sections that do not apply.

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION

1a.4.1. Guideline citation (*including date*) and **URL for guideline** (*if available online*):

1a.4.2. Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

1a.4.3. Grade assigned to the quoted recommendation with definition of the grade:

1a.4.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: If separate grades for the strength of the evidence, report them in section 1a.7.*)

1a.4.5. Citation and URL for methodology for grading recommendations (*if different from 1a.4.1*)**:**

ICSI GRADE Methodology: https://www.icsi.org/_asset/7mtqyr/ReviewingEvidenceUsingGRADE.pdf

1a.4.6. If guideline is evidence-based (rather than expert opinion), are the details of the quantity, quality, and consistency of the body of evidence available (e.g., evidence tables)?

☐ Yes → ***complete section 1a.7***

☐ No → ***report on another systematic review of the evidence in sections 1a.6 and 1a.7; if another review does not exist, provide what is known from the guideline review of evidence in 1a.7***

1a.5. UNITED STATES PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

1a.5.1. Recommendation citation (*including date*) and **URL for recommendation** (*if available online*):

1a.5.2. Identify recommendation number and/or page number and **quote verbatim, the specific recommendation**.

1a.5.3. Grade assigned to the quoted recommendation with definition of the grade:

1a.5.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: the grading system for the evidence should be reported in section 1a.7.*)

1a.5.5. Citation and URL for methodology for grading recommendations (*if different from 1a.5.1*)**:**

Complete section 1a.7

1a.6. OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1. Citation (*including date*) and **URL** (*if available online*):

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

Complete section 1a.7

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

If more than one systematic review of the evidence is identified above, you may choose to summarize the one (or more) for which the best information is available to provide a summary of the quantity, quality, and consistency of the body of evidence. Be sure to identify which review is the basis of the responses in this section and if more than one, provide a separate response for each review.

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

1a.7.2. Grade assigned for the quality of the quoted evidence with definition of the grade:

1a.7.3. Provide all other grades and associated definitions for strength of the evidence in the grading system.

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range: 1998 to 2013

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

1a.7.6. What is the overall quality of evidence across studies in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) across studies in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for each new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

1a.8 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.8.1 What process was used to identify the evidence?

1a.8.2. Provide the citation and summary for each piece of evidence.

Component # 4 Daily Aspirin or Anti-platelet Medication

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)
Outcome

- ☐ Health outcome: Click here to name the health outcome
- ☐Patient-reported outcome (PRO): Click here to name the PRO
  PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors
- ☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☒ Process: Appropriate use of daily aspirin or anti-platelet medication (if patient with ischemic vascular disease is without contraindications or exceptions to aspirin or anti-platelet use)

☐ Structure: Click here to name the structure

☐ Other: Click here to name what is being measured

HEALTH OUTCOME/PRO PERFORMANCE MEASURE *If not a health outcome or PRO, skip to 1a.3*

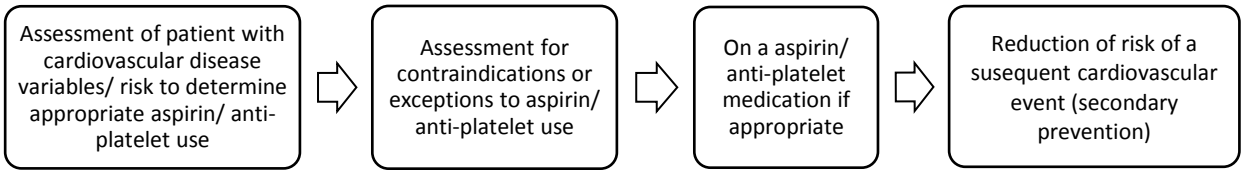
1a.2. Briefly state or diagram the path between the health outcome (or PRO) and the healthcare structures, processes, interventions, or services that influence it.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process, intervention, or service (*i.e., influence on outcome/PRO*).

Note: For health outcome/PRO performance measures, no further information is required; however, you may provide evidence for any of the structures, processes, interventions, or service identified above.

INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

1a.3. Briefly state or diagram the path between structure, process, intermediate outcome, and health outcomes. Include all the steps between the measure focus and the health outcome.



1a.3.1. What is the source of the systematic review of the body of evidence that supports the performance measure?

☒ Clinical Practice Guideline recommendation – ***complete sections 1a.4, and 1a.7***

☐ US Preventive Services Task Force Recommendation – ***complete sections 1a.5 and 1a.7***

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*) – ***complete sections 1a.6 and 1a.7***

☐ Other – ***complete section 1a.8***

Please complete the sections indicated above for the source of evidence. You may skip the sections that do not apply.

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION

1a.4.1. Guideline citation (*including date*) and **URL for guideline** (*if available online*):

ICSI Stable Coronary Artery Disease April 2011
http://www.icsi.org/guidelines_and_more/gl_os_prot/cardiovascular/coronary_artery_disease/coronary_artery_disease__stable__3.html

AHA/ACCF Secondary Prevention and Risk Reduction Therapy for Patients with Coronary and Other Atherosclerotic Vascular Disease: 2011 Update http://circ.ahajournals.org/content/early/2011/11/01/CIR.0b013e318235eb4d.full.pdf

1a.4.2. Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

ICSI Stable Coronary Artery Disease April 2011

Address Modifiable Risk Factors and Comorbid Conditions:

Comorbid conditions that could affect myocardial ischemia may include hypertension, anemia, thyroid disease, hypoxemia and others. Modifiable risk factors for coronary heart disease need to be evaluated and may include smoking, inadequate physical activity, stress, hyperlipidemia, obesity, hypertension and diabetes mellitus.

Intervention involving any risk factor pertinent to the patient is encouraged and may include education, goal setting, and follow-up as necessary (Rutherford, 1992 [R]; Shub, 1990 [R]).

Antiplatelet Therapy:

The use of one aspirin tablet daily (81-162 mg) is strongly recommended unless there are medical contrain-dications (Antiplatelet Trialists´ Collaboration, 1994 [A]; CAPRI, 1996 [A]; Fuster, 1993 [R]; Juul-Möller, 1992 [A]; Kurth, 2003 [A]; Ridker, 1991 [A]).The Antithrombotic Trialists´ Collaboration is a meta-analysis that analyzed 287 studies involving 135,000 patients for different aspects of antiplatelet therapy. When comparing the 500-1,500 mg versus 160-325 mg versus 75-150 mg daily regimens of aspirin in multiple trials, there was a trend of reduction in vascular events with decreased dose (odds reduction: 19% versus 26% versus 32%, respectively) (Antithrombotic Trialists Collaboration; 2002 [M]). Although the meta-analysis concludes that risk of gastrointestinal bleed was similar among doses 325 mg or less, other studies such as the CURE study showed increased bleeding risk with increasing the dose, without any increase in efficacy (Peters, 2003 [A]).

The authors conclude that aspirin dose in the range of 75-150 mg should be given for the long-term prevention of serious vascular events in high risk patients, and that there may be a reduced benefit when increasing the dose over 150 mg daily. Doses available to most clinicians are in increments of 81 mg; therefore, the recommended dose range is 81-162 mg daily.

AHA/ACCF Secondary Prevention and Risk Reduction Therapy for Patients with Coronary and Other Atherosclerotic Vascular Disease: 2011 Update

Antiplatelet agents/anticoagulants

(pg. 3 – 4; recommendations 1 and 5 apply for patients with ischemic vascular disease)
Class I Recommendations:
1. Aspirin 75–162 mg daily is recommended in all patients with coronary artery disease unless contraindicated. (Level of Evidence: A)
- Clopidogrel 75 mg daily is recommended as an alternative for patients who are intolerant of or allergic to aspirin. (Level of Evidence: B)

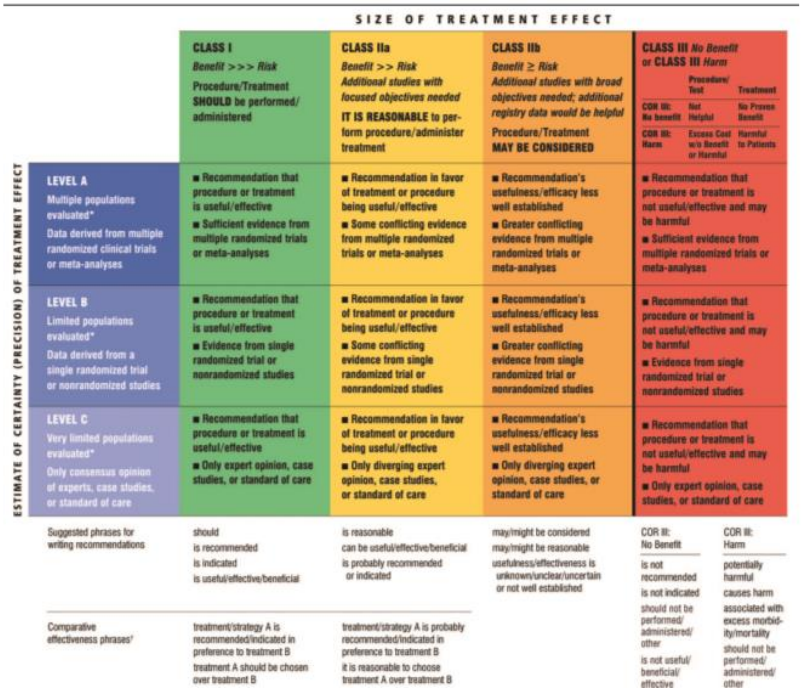
5. For patients with symptomatic atherosclerotic peripheral artery disease of the lower extremity, antiplatelet therapy with aspirin (75–325 mg daily) or clopidogrel (75 mg daily) should be started and continued. (Level of Evidence: A)

1a.4.3. Grade assigned to the quoted recommendation <u>with definition</u> of the grade:

Class I = Benefit >>> outweighs risk. Procedure/ Treatment SHOULD be performed/ administered
Level A = Multiple populations evaluated. Data derived from multiple randomized clinical trials or meta-analyses. Recommendation that procedure or treatment is useful/ effective. Sufficient evidence from multiple randomized control trials or meta-analysis.
Level B = Limited populations evaluated. Data derived from a single randomized trial or non-randomized studies

1a.4.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: If separate grades for the strength of the evidence, report them in section 1a.7.*)

Table 2. Applying Classification of Recommendation and Level of Evidence



ESTIMATE OF CERTAINTY (PRECISION) OF TREATMENT EFFECT	SIZE OF TREATMENT EFFECT: CLASS I *Benefit >>> Risk* Procedure/Treatment **SHOULD** be performed/ administered	CLASS IIa *Benefit >> Risk* *Additional studies with focused objectives needed* **IT IS REASONABLE** to perform procedure/administer treatment	CLASS IIb *Benefit ≥ Risk* *Additional studies with broad objectives needed; additional registry data would be helpful* Procedure/Treatment **MAY BE CONSIDERED**	CLASS III *No Benefit* or CLASS III *Harm* Procedure/ Test / Treatment COR III: No benefit — Not Helpful — No Proven Benefit COR III: Harm — Excess Cost w/o Benefit or Harmful — Harmful to Patients
LEVEL A Multiple populations evaluated* Data derived from multiple randomized clinical trials or meta-analyses	■ Recommendation that procedure or treatment is useful/effective ■ Sufficient evidence from multiple randomized trials or meta-analyses	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Sufficient evidence from multiple randomized trials or meta-analyses
LEVEL B Limited populations evaluated* Data derived from a single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is useful/effective ■ Evidence from single randomized trial or nonrandomized studies	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Evidence from single randomized trial or nonrandomized studies
LEVEL C Very limited populations evaluated* Only consensus opinion of experts, case studies, or standard of care	■ Recommendation that procedure or treatment is useful/effective ■ Only expert opinion, case studies, or standard of care	■ Recommendation in favor of treatment or procedure being useful/effective ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation's usefulness/efficacy less well established ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Only expert opinion, case studies, or standard of care
Suggested phrases for writing recommendations	should is recommended is indicated is useful/effective/beneficial	is reasonable can be useful/effective/beneficial is probably recommended or indicated	may/might be considered may/might be reasonable usefulness/effectiveness is unknown/unclear/uncertain or not well established	COR III: No Benefit is not recommended is not indicated should not be performed/ administered/ other is not useful/ beneficial/ effective COR III: Harm potentially harmful causes harm associated with excess morbidity/mortality should not be performed/ administered/ other
Comparative effectiveness phrases†	treatment/strategy A is recommended/indicated in preference to treatment B treatment A should be chosen over treatment B	treatment/strategy A is probably recommended/indicated in preference to treatment B it is reasonable to choose treatment A over treatment B		

1a.4.5. Citation and URL for methodology for grading recommendations (*if different from 1a.4.1*)**:**

1a.4.6. If guideline is evidence-based (rather than expert opinion), are the details of the quantity, quality, and consistency of the body of evidence available (e.g., evidence tables)?

☒ Yes → ***complete section 1a.7***

☐ No → ***report on another systematic review of the evidence in sections 1a.6 and 1a.7; if another review does not exist, provide what is known from the guideline review of evidence in 1a.7***

1a.5. UNITED STATES PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

1a.5.1. Recommendation citation (*including date*) and **URL for recommendation** (*if available online*):

1a.5.2. Identify recommendation number and/or page number and **quote verbatim, the specific recommendation**.

1a.5.3. Grade assigned to the quoted recommendation with definition of the grade:

1a.5.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: the grading system for the evidence should be reported in section 1a.7.*)

1a.5.5. Citation and URL for methodology for grading recommendations (*if different from 1a.5.1*)**:**

Complete section 1a.7

1a.6. OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1. Citation (*including date*) and **URL** (*if available online*):

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

Complete section 1a.7

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

If more than one systematic review of the evidence is identified above, you may choose to summarize the one (or more) for which the best information is available to provide a summary of the quantity, quality, and consistency of the body of evidence. Be sure to identify which review is the basis of the responses in this section and if more than one, provide a separate response for each review.

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

AHA/ACCF Secondary Prevention and Risk Reduction Therapy for Patients with Coronary and Other Atherosclerotic Vascular Disease: 2011 Update

The prevention of secondary cardiovascular events for patients with cardiovascular disease by the appropriate prescribing of aspirin or anti-platelet medications.

1a.7.2. Grade assigned for the quality of the quoted evidence <u>with definition</u> of the grade:

Class I = Benefit >>> outweighs risk. Procedure/ Treatment SHOULD be performed/ administered

Level A = Multiple populations evaluated. Data derived from multiple randomized clinical trials or meta-analyses. Recommendation that procedure or treatment is useful/ effective. Sufficient evidence from multiple randomized control trials or meta-analysis.

1a.7.3. Provide all other grades and associated definitions for strength of the evidence in the grading system.

Table 2. Applying Classification of Recommendation and Level of Evidence

ESTIMATE OF CERTAINTY (PRECISION) OF TREATMENT EFFECT	SIZE OF TREATMENT EFFECT: CLASS I *Benefit >>> Risk* Procedure/Treatment **SHOULD** be performed/ administered	CLASS IIa *Benefit >> Risk* *Additional studies with focused objectives needed* **IT IS REASONABLE** to perform procedure/administer treatment	CLASS IIb *Benefit ≥ Risk* *Additional studies with broad objectives needed; additional registry data would be helpful* Procedure/Treatment **MAY BE CONSIDERED**	CLASS III *No Benefit* or CLASS III *Harm* Procedure/Test: COR III: No benefit — Not Helpful; COR III: Harm — Excess Cost w/o Benefit or Harmful Treatment: No Proven Benefit; Harmful to Patients
LEVEL A Multiple populations evaluated* Data derived from multiple randomized clinical trials or meta-analyses	■ Recommendation that procedure or treatment is useful/effective ■ Sufficient evidence from multiple randomized trials or meta-analyses	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Sufficient evidence from multiple randomized trials or meta-analyses
LEVEL B Limited populations evaluated* Data derived from a single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is useful/effective ■ Evidence from single randomized trial or nonrandomized studies	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Evidence from single randomized trial or nonrandomized studies
LEVEL C Very limited populations evaluated* Only consensus opinion of experts, case studies, or standard of care	■ Recommendation that procedure or treatment is useful/effective ■ Only expert opinion, case studies, or standard of care	■ Recommendation in favor of treatment or procedure being useful/effective ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation's usefulness/efficacy less well established ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Only expert opinion, case studies, or standard of care
Suggested phrases for writing recommendations	should is recommended is indicated is useful/effective/beneficial	is reasonable can be useful/effective/beneficial is probably recommended or indicated	may/might be considered may/might be reasonable usefulness/effectiveness is unknown/unclear/uncertain or not well established	COR III: No Benefit: is not recommended; is not indicated; should not be performed/ administered/ other; is not useful/ beneficial/ effective COR III: Harm: potentially harmful; causes harm; associated with excess morbidity/mortality; should not be performed/ administered/ other
Comparative effectiveness phrases†	treatment/strategy A is recommended/indicated in preference to treatment B; treatment A should be chosen over treatment B	treatment/strategy A is probably recommended/indicated in preference to treatment B; it is reasonable to choose treatment A over treatment B		

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range: 2002 to 2009

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

One meta-analysis of 22 randomized control trials.

One collaborative meta-analysis involving 287 studies, 135,000 patients therapy vs. control and 77,000 patients comparing different anti-platelet regimens.

1a.7.6. What is the overall quality of evidence <u>across studies</u> in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

Since the 2006 update of the American Heart Association (AHA)/American College of Cardiology Foundation (ACCF) guidelines on secondary prevention, important evidence from clinical trials has emerged that further supports and broadens the merits of intensive risk-reduction therapies for patients with established coronary and other atherosclerotic vascular disease, including peripheral artery disease, atherosclerotic aortic disease, and carotid artery disease. In reviewing this evidence and its clinical impact, the writing group believed it would be more appropriate to expand the title of this guideline to "Secondary Prevention and Risk Reduction Therapy for Patients with Coronary and Other Atherosclerotic Vascular Disease." Indeed, the growing body of evidence confirms that in patients with atherosclerotic vascular disease, comprehensive risk factor management reduces risk as assessed by a variety of outcomes, including improved survival, reduced recurrent events, the need for revascularization procedures, and improved quality of life. It is important not only that the healthcare provider implement these recommendations in appropriate patients but also that healthcare systems support this implementation to maximize the benefit to the patient.

The quality of evidence across studies related to the secondary prevention of cardiovascular events was rated as strong. [Level A = Multiple populations evaluated. Data derived from multiple randomized clinical trials or meta-analyses. Recommendation that procedure or treatment is useful/ effective. Sufficient evidence from multiple randomized control trials or meta-analysis.]

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) <u>across studies</u> in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

Please see conclusions from the guideline group in section 1a.7.6.

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

Please see conclusions from the guideline group in section 1a.7.6.

The estimate of benefit outweighs significantly with a Class I Level A grade recommendation.

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for <u>each</u> new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

1a.8 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.8.1 What process was used to identify the evidence?

1a.8.2. Provide the citation and summary for each piece of evidence.

1. Evidence, Performance Gap, Priority – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all subcriteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form
0076_OVC_Template_Evidence_MNCM_2016.docx

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- disparities in care across population groups.

1b.1. **Briefly explain the rationale for this measure** *(e.g., the benefits or improvements in quality envisioned by use of this measure)*
The intermediate physiological and biochemical outcomes included in this composite measure along with the appropriate use of statins and daily aspirin or anti-platelets are modifiable lifestyle risk factors that can ultimately decrease the incidence of long term catastrophic events and chronic illness associated with cardiovascular disease.

1b.2. **Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.** *(This is required for endorsement maintenance. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included). This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*
For 2016 (2015 dates of service), 66.1% of the patients met all four component targets in the composite measure and were considered optimally managed. This rate is a weighted average of the total population of patients for clinics submitting data (Total Population = 104,494, Submitted = 104,395). 99.5% of the clinics submitted full population data; the remaining clinics provided a random sample. Of the clinics that were reportable (patient n >= 30), there was a wide range of variability with the lowest scoring clinic at 8.6% and the highest scoring clinic at 85.3%.

The trends for this measure are as follows:

Report Year	Rate	Patients (Den)	Numerator	Eligible	% submit/eligible
2007	38.9%	4,662	1,595	11,740	39.7%
2008	32.6%	36,126	11,997	54,708	66.0%
2009	33.8%	46,779	16,529	80,907	57.8%
2010	33.8%	63,241	21,589	95,791	66.0%
2011*	39.7%	66,910	27,083	96,270	69.5%
2012	49.4%	78,886	39,242	95,482	82.6%
2013	48.5%	87,345	42,689	93,761	93.1%
2014	50.0%	98,803	49,408	99,550	99.2%
2015**	69.3%	102,654	71,196	103,006	99.7%
2016***	66.1%	104,395	69,026	104,494	99.9%

* Blood pressure component target change based on evidence/ guidelines from < 130/80 to < 140/90
** Cholesterol management component suppressed during re-design
*** Cholesterol management component change from LDL < 100 to appropriate statin use

Individual rates of the components are as follows:
Blood Pressure <140/90 = 85.0%
Statin Use = 95.2%
Daily Aspirin Use = 96.7%
Tobacco Non-user = 83.0%

Please note that while the all-or-none composite measure is considered to be the gold standard, reflecting best patient outcomes, the individual components may be measured as well. This is particularly helpful in quality improvement efforts to better understand where opportunities exist in moving the patients toward achieving all of the desired outcomes. Please refer to the additional

numerator logic provided for each component.

Trend over time by Component and Report Year

	2009	2010	2011	2012	2013	2014	2015	2016
BP <140/90	-	-	-	84.0%	84.1%	84.9%	85.2%	85.0%
Aspirin Use	92.5%	91.9%	94.2%	94.7%	96.5%	96.6%	96.6%	96.7%
Tobacco Free	82.4%	81.2%	82.7%	82.6%	82.9%	84.1%	83.5%	83.0%

Weighted Mean: 66.1%
Mean: 63.9%
Median: 65.8%
Standard Deviation: 0.10564 (10.6%)
Min: 8.6%
Max: 85.3%
(reflects reportable clinics, patient n >= 30)

Distribution of Rates Among Clinics: (reportable clinics)

0%-9.9%	0.2%
10%-19.9%	0.0%
20%-29.9%	0.0%
30%-39.9%	2.2%
40%-49.9%	8.6%
50%-59.9%	17.5%
60%-69.9%	38.9%
70%-79.9%	30.3%
80%-89.9%	2.2%

Clinic Rates by Decile:

Percentile	Rate
10th	48.8 %
20th	56.3 %
30th	60.2 %
40th	63.4 %
50th	65.8 %
60th	68.1 %
70th	70.3 %
80th	72.4 %
90th	75.6 %

Consumer facing website MN HealthScores
Displays the top 10 best performers (2015 Report/ 2014 Dates of Service) in addition to rates for all clinics in MN
http://www.mnhealthscores.org/vascular-care

Rate %	Clinic Location	
90%	Allina Health- Hopkins (Aspen)	St Paul, MN
88%	Allina Health- Vadnais Heights	Vadnais Heights, MN
86%	Fairview Rosemount Clinic	Rosemount, MN
85%	Entira Family Clinics- Banning Ave	White Bear Lake, MN
85%	Park Nicollet Clinic- Carlson	Minnetonka, MN
84%	Park Nicollet Clinic- Prairie Center	Eden Prairie, MN
84%	Park Nicollet Clinic - Eagan	Eagan, MN
84%	Mankato Clinic - Wickersham	Mankato, MN
83%	Sanford Health Vermillion Clinic	Vermillion, MN
83%	Allina Health - Woodbury	Woodbury, MN

Publicly reported data with clinic level rates is available on the MN HealthScores website at
http://www.mnhealthscores.org/medical-group-measure-detail/vascular-disease-0/#/results

In 2016 (2015 dates of service), 111 medical groups representing 671 physician clinics and 104,494 patients with IVD in Minnesota and neighboring communities submitted data for this measure. Of the 104,494 IVD patients, 104,395 patients were submitted for rate calculation. 99.6% of the clinics submitted full population data, with only 3 clinics submitting a random sample. Dates of service included 01/01/2015 to 12/31/2015.

The data submitted represents 99.9% of all eligible patients; based on the large sample size, the results can be reliably reproduced. The data submission process requires individual patient data for each component of the "all or none" composite measure (e.g., most recent blood pressure values in the measurement period). This information is accurately captured as evidenced by post submission validation audits against the patient's medical record.

Characteristics of the entities reporting data:
Based on number of physicians, the size of the 111 medical groups that submitted data ranged from one-physician practices to medical groups with more than 500 physicians . Ranges include: Medical groups with <25 physicians = 80; medical groups with 25-99 physicians = 15; medical groups with 100-249 physicians = 9; medical groups with 250+ physicians = 7. 39 medical groups were located within the Twin Cities metro area, while 72 medical groups were located outside of the Twin Cities metro area. 52 medical groups were identified as primary care clinics, 52 medical groups were identified as multi-specialty clinics with 22 inclusive of cardiology, and one group was identified as a single-specialty clinic (cardiology).

Of the 671 clinic sites that reported data, 57% used their EMR exclusively to extract data, 40% supplemented data extraction with some chart abstraction and 3% relied on manual abstraction processes of their EMR or paper based record. Interestingly, in 2010 118 clinics in MN had paper based system, in 2015 only 5 clinics are on a paper based system.

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for endorsement maintenance. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*

Race	Optimal Rate*	Num/Den	Rate (2014)**	Change
White	67.2%	(59,562/88,631)	50.8%	16.4
Black or African American	47.6%	(1,451/3,049)	35.5%	12.1
Asian	70.6%	(1,098/1,554)	54.4%	16.2
Multi-Racial	53.4%	(234/359)	42.6%	10.8
Chose not to disclose	68.5%	(650/949)	54.3%	14.2
American Indian/Alaska Native	51.8%	(5,449/1,059)	34.6%	17.2
Some Other Race	70.4%	(178/253)	55.1%	15.3
Native Hawaiian/Pacific Isl	71.4%	(70/98)	50.0%	21.4
Unknown	61.3%	(19/31)	43.6%	17.7
Race not Reported	62.0%	(5,215/8,412)	48.9%	13.1
Grand Total	66.1%	69,026/104,395)	50.0%	16.1

Gender	Optimal Rate*	Num/Den	Rate (2014)**	Change
F	62.5%	(18,809/30,116)	44.8%	17.7
M	67.6%	(24,062/74,279)	52.1%	15.5
Grand Total	66.1%	(69,026/104,395)	50.0%	16.1

Age Band	Optimal Rate*	Num/Den	Rate (2014)**	Change
18 to 25	37.0%	(20/54)	24.4%	12.6
26 to 50	53.5%	(3,681/6,880)	35.9%	17.6
51 to 65	62.6%	(27,902/44,563)	46.9%	15.7
66 to 75	70.7%	(37,423/52,898)	54.8%	15.9
Grand Total	66.1%	(69,026/104,395)	50.0%	16.1

* Cholesterol management component redesigned to appropriate statin use
** Cholesterol management component was LDL < 100

Heart disease is the leading cause of death for people of most racial/ethnic groups in the United States, including African Americans, Hispanics, and whites. For Asian Americans or Pacific Islanders and American Indians or Alaska Natives, heart disease is second only to cancer. [Heron M. Deaths: Leading causes for 2008[PDF-2.7M]. National vital statistics reports. 2012;60(6)]

Cardiovascular disease risk factors include both modifiable risk factors (hypocholesteremia, hypertension, diabetes and pre-diabetes, overweight and obesity, tobacco use, lack of physical activity, unhealthy diet and stress) and risk factors that cannot be changed (age, gender and family history of CHD). [National Heart, Lung, Blood Institute www.nhlbi.nih.gov/health/health-topics/topics/hd]

Blacks are nearly twice as likely to have a first stroke and much more likely to die from one than whites. [Heart Disease and Stroke Statistics – 2009 Update: A Report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Circulation 2008]

American Indians/Alaska Natives die from heart disease much earlier than expected – 36% are under 65 compared with only 17% for the U.S. population overall. [Centers for Disease Control and Prevention. Disparities in premature deaths from heart disease, 2001. MMWR 53(6):121-125]

MNCM publishes an annual Health Equity of Care Report stratifying measures by race Hispanic ethnicity, primary language and country of origin. http://mncm.org/wp-content/uploads/2015/11/2015-Health-Equity-of-Care-Report-Final-2.11.2016.pdf.
For the Optimal Vascular Care measure:

Statewide rates for each racial group were stratified by region for comparative analyses. The highest and lowest rates for each racial group were found in the following regions:
* Race – For almost all racial groups, their highest rate was found in the East Metro and West Metro regions. The Asian racial group had the highest rate in four of the five regions where the group was reportable. The White population had the highest rate in four regions and was the only reportable racial group in two regions (Southeast and Southwest). The Black or African American racial group had the lowest rate in four regions.
* Hispanic Ethnicity – Hispanics had the highest rate in five regions. Hispanics and Non-Hispanics had the same rate in the Central region. Non-Hispanics had a higher rate in the West Metro region.
* Preferred Language – In four regions, the only reportable preferred language group was English and this group's highest rate was found in the West Metro region. Patients who preferred speaking Spanish had the highest rate in two regions.
* Country of Origin – The United States was the only reportable country of origin group in three regions. This group's highest rate was found in the West Metro region and it had the lowest rate in three regions. Patients born in an Asian country had the highest rate in three regions.

American Indian or Alaskan Native: Highest rate East Metro and West Metro regions at 66%; lowest rate Central region at 36%.
Asian: Highest rate East Metro region at 83%; lowest rate Central region at 70%.
Black or African American: Highest rate East Metro region at 64%; lowest rate Minneapolis region at 47%.
Multi-Racial: Highest rate East Metro region at 76%; lowest rate West Metro region at 66%.
Native Hawaiian or Other Pacific Islander: did not meet the minimum reporting threshold
Some Other Race: Highest rate West Metro region at 78%; lowest rate East Metro region at 65%.
White: Highest rate West Metro region at 75%; lowest rate Northeast and Southwest regions at 66%.

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations.

1c. High Priority (previously referred to as High Impact)

The measure addresses:

- a specific national health goal/priority identified by DHHS or the National Priorities Partnership convened by NQF; OR
- a demonstrated high-priority (high-impact) aspect of healthcare (e.g., affects large numbers of patients and/or has a substantial impact for a smaller population; leading cause of morbidity/mortality; high resource use (current and/or future); severity of illness; and severity of patient/societal consequences of poor quality).

1c.1. **Demonstrated high priority aspect of healthcare**
Affects large numbers, A leading cause of morbidity/mortality, Severity of illness
1c.2. **If Other:**

1c.3. **Provide epidemiologic or resource use data that demonstrates the measure addresses a high priority aspect of healthcare. List citations in 1c.4.**
According to the MN Department of Health, cardiovascular disease is a high impact clinical condition in Minnesota.
In 2014, 3.8% of adults in Minnesota reported ever having had a heart attack in their lifetime – over 150,000 people. More than 18% of all deaths in Minnesota are due to heart disease (7,571 deaths in 2014), making it the 2nd-leading cause of death in the state behind cancer.
In 2013, Minnesotans experienced almost 45,000 acute heart disease hospitalizations and Minnesota had the lowest overall heart disease mortality rate in the United States.

1c.4. **Citations for data demonstrating high priority provided in 1a.3**
Heart disease in Minnesota http://www.health.state.mn.us/divs/healthimprovement/data/quick-facts/heartdisease.html
[Minnesota Behavioral Risk Factor Surveillance System]
[Vital Statistics 2009-2014. MN Center for Health Statistics, MDH]
[MN Hospital Uniform Billing (UB) Claims Data, Health Economics Program, MDH and Minnesota Hospital Association.]
[CDC, National Center for Health Statistics, Compressed Mortality File (CMF) on CDC WONDER Online Database]

1c.5. **If a PRO-PM (e.g. HRQoL/functional status, symptom/burden, experience with care, health-related behaviors), provide evidence that the target population values the measured PRO and finds it meaningful.** *(Describe how and from whom their input was obtained.)*

1d. Composite Quality Construct and Rationale

1d.1. **A composite performance measure is a combination of two or more component measures, each of which individually reflects quality of care, into a single performance measure with a single score.**

For purposes of NQF measure submission, evaluation, and endorsement, the following will be considered composites:

- Measures with two or more individual performance measure scores combined into one score for an accountable entity.
- Measures with two or more individual component measures assessed separately for each patient and then aggregated into one score for an accountable entity:
    - all-or-none measures (e.g., all essential care processes received, or outcomes experienced, by each patient); or
    - any-or-none measures (e.g., any or none of a list of adverse outcomes experienced, or inappropriate or unnecessary care processes received, by each patient).

1d.1. Please identify the composite measure construction: all-or-none measures (e.g., all essential care processes received, or outcomes experienced, by each patient)

1d.2. **Describe the quality construct, including:**

- the overall area of quality
- included component measures and
- the relationship of the component measures to the overall composite and to each other.

This composite measure is a patient level all-or-none composite in which the desired goal is for the patient is to achieve multiple intermediate physiological clinical outcome and medication use targets to best reduce their overall risk of developing further ischemic vascular complications (short and long term) or an additional cardiovascular event. Reducing modifiable risks was the reason why this measure was developed. The components of this measure include blood pressure control, appropriate use of statins, appropriate use of daily aspirin or anti-platelet medication and being tobacco-free.

1. Blood pressure less than 140/90 mmHg
2. On a statin medication, unless allowed contraindications or exceptions are present
3. Non-tobacco user
4. On daily aspirin or anti-platelet medication, unless allowed contraindications or exceptions are present

Numerator is calculated at the patient level and numerator compliance is defined as the patient achieving all four components of the

measure. The components are weighted equally.

1d.3. **Describe the rationale for constructing a composite measure, including how the composite provides a distinctive or additive value over the component measures individually.**
Achieving the intermediate physiological outcome targets related to blood pressure in addition to being tobacco free and use of daily aspirin and statins where appropriate are the cardiovascular patient's best mechanisms of avoiding or postponing long term complications associated with this chronic condition which affects millions of Americans. Measuring providers separately on individual targets is not as patient centric as a measure that seeks to reduce multiple risk factors for each patient. Patients with ischemic vascular disease are more likely to reduce their overall risk and maximize health outcomes by achieving several intermediate physiological targets and medication use targets.

Please note that while the all-or-none composite measure is considered to be the gold standard, reflecting best patient outcomes, the individual components may be measured as well. This is particularly helpful in quality improvement efforts to better understand where opportunities exist in moving the patients toward achieving all of the desired outcomes. Please refer to the additional numerator logic provided for each component.

1d.4. **Describe how the aggregation and weighting of the component measures are consistent with the stated quality construct and rationale.**
Numerator is calculated at the patient level and numerator compliance is defined as the patient achieving all four components of the measure. The components are treated equally; there is no weighting. Some of the components have an exception methodology within allowing a "free-pass" on the component if it does not apply to the patient.

Most recent blood pressure in the measurement period is less than 140 systolic AND less than 90 diastolic (applies to all denominator patients)
AND
Statin Use if appropriate and no contraindications/ exceptions
IVD patients age 18-20 "free-pass"
Age 21 to 75 on statin unless LDL < 40 or contraindications/exceptions
AND
Patient's tobacco status (documented within the last 2 years) is tobacco free (applies to all denominator patients)
AND
Daily aspirin or anti-platelet use unless contraindications/exceptions.

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the subcriteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. **Subject/Topic Area** *(check all the areas that apply):*
Cardiovascular, Cardiovascular : Ischemic Heart Disease, Coronary Artery Disease

De.6. **Cross Cutting Areas** *(check all the areas that apply):*
Patient and Family Engagement

S.1. **Measure-specific Web Page** *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*
http://mncm.org/cycle-a-dds-guides/

S.2a. **If this is an eMeasure**, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. **Data Dictionary, Code Table, or Value Sets** *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*
Attachment **Attachment:** MNCM_-0076_Optimal_Vascular_Care_Specs_Fields_RA_2-2016.xlsx

S.3. **For endorsement maintenance**, please briefly describe any changes to the measure specifications since last endorsement date and explain the reasons.
The cholesterol component of this measure was re-designed in 2014 as a result of significant changes to guidelines for cholesterol management released November 2013 by the American College of Cardiology and the American Heart Association. Previously the cholesterol component was an intermediate outcome defined as target LDL < 100, which is no longer supported by evidence and guidelines. The redesigned cholesterol component was completed by the measure development work group in October of 2014 and focuses on appropriate statin use. This change in the measure will be effective for the 2016 reporting year for 2015 dates of service (1/1/2015 to 12/31/2015).

S.4. **Numerator Statement** *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome)*
IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
The number of patients in the denominator whose IVD was optimally managed during the measurement period as defined by achieving ALL of the following:
• The most recent blood pressure in the measurement period has a systolic value of less than 140 mmHg AND a diastolic value of less than 90 mmHg
• On a statin medication, unless allowed contraindications or exceptions are present
• Patient is not a tobacco user
• On daily aspirin or anti-platelet medication, unless allowed contraindications or exceptions are present

S.5. **Time Period for Data** *(What is the time period in which data will be aggregated for the measure, e.g., 12 mo, 3 years, look back to August for flu vaccination? Note if there are different time periods for the numerator and denominator.)*
Values are collected as the most recent during the measurement period (January 1 through December 31).

S.6. **Numerator Details** *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
In order to be numerator compliant all four components must be met
* Blood pressure less than 140/90 mmHg AND
* On a statin medication, unless allowed contraindications or exceptions are present AND
* Non-tobacco user AND
* On daily aspirin or anti-platelet medication, unless allowed contraindications or exceptions are present

BLOOD PRESSURE COMPONENT
Blood Pressure Date [Date (mm/dd/yyyy)] AND
BP Systolic [Numeric] AND
BP Diastolic [Numeric]
Numerator component calculation: numerator component compliant is BP during the measurement year AND Systolic < 140 AND Diastolic < 90.

BP Date
Enter the date of the most recent blood pressure result during the measurement period.
* Do NOT enter a test date that occurred in yyyy (beyond measurement period). A date in yyyy will create an ERROR upon submission.
* A test result from a provider outside of the reporting medical group is allowed if the result is documented in the reporting medical group's patient record and is the most recent test result during the measurement period.
* Do NOT enter a blood pressure result that is reported by or taken by the patient.
* Do NOT enter a blood pressure result obtained in the following care settings: Inpatient, Emergency Department, Urgent Care,

other settings designated for surgical or diagnostic procedures or an office visit associated with acute pain or pain of at least moderate severity (greater than or equal to four on a scale of zero to 10).

BP Systolic
Enter the value of the most recent systolic blood pressure result during the measurement period.
* If more than one value is recorded on the most recent date, the lowest value may be submitted. It does NOT need to be from the same reading.
NOTE: The systolic blood pressure is the upper number in the recorded fraction. For example, the systolic value for a blood pressure of 124/72 mmHg is 124.

BP Diastolic
Enter the value of the most recent diastolic blood pressure result during the measurement period.
* If more than one value is recorded on the most recent date, the lowest value may be submitted. It does NOT need to be from the same reading.
NOTE: The diastolic blood pressure is the lower number in the recorded fraction. For example, the diastolic value for a blood pressure of 124/72 mmHg is 72.

Leave BLANK if a blood pressure was not obtained during the measurement period.

CHOLESTEROL MANAGEMENT STATIN COMPONENT
LDL Date [Date (mm/dd/yyyy)] AND
LDL Value [Numeric]
For calculating exceptions to statin use based on very low LDL (< 40 for cardiovascular disease and < 70 for patients with diabetes)

Enter the date of the most recent LDL test result between mm/dd/yyyy and mm/dd/yyyy (five year range including measurement period)
* Do NOT enter a test date that occurred in yyyy. A date in yyyy (beyond measurement period) will create an ERROR upon submission.
* A test result from a provider outside of the reporting medical group is allowed if the result is documented in the reporting medical group's patient record and is the most recent test result within the allowable time period.
* If the LDL result is too high to calculate, still enter the LDL test date if it is the most recent test result within the allowable time period.
LDL values within the last five years will be used to calculate potential exceptions to being on a statin medication. Leave BLANK if an LDL test was not performed between mm/dd/yyyy and mm/dd/yyyy (five year range including measurement period).

Statin Medication [Numeric] AND
Statin Medication Date [Date (mm/dd/yyyy)] AND/OR
Station Medication Exception [Numeric] AND
Station Medication Exception Date [Date (mm/dd/yyyy)]
Numerator component calculation: numerator component compliant if on a statin (prescribed/ ordered) or low LDL value (see above) or documented contraindication/exception is present.

Statin Medication:
Enter the code that corresponds to whether the patient was prescribed a statin medication or if a statin medication was active on the patient's medication list at any time during the measurement period.
Please see Appendix C for a list of statin medications.
1 = Yes, patient was prescribed a statin medication or a statin medication was reviewed and active on the patient's medication list.
2 = No, patient was not prescribed a statin medication and a statin medication was not reviewed and active on the patient's medication list.
The following exception to statin medication use will be identified by the portal based on the submitted LDL values
* Patients aged 21 to 75 years and an LDL result less than 40 mg/dL
A blank field will create an ERROR upon submission.

Statin Medication Date
Enter the date of the most recent statin prescription, order or review on an active medications list that included a statin during the measurement period.
* Do NOT enter a date that occurred in yyyy. A date in yyyy (beyond measurement period) will create an ERROR upon submission.

* If a statin was not prescribed, ordered, or reviewed as an active medication during the measurement period, leave BLANK.

Station Medication Exception
If the patient was NOT prescribed a statin medication during the measurement period (Field AA = 2), enter the value that corresponds to any of the following contraindications or exceptions:
1 = Pregnancy at any time during the measurement period
2 = Active liver disease (liver failure, cirrhosis, hepatitis)
3 = Rhabdomyolysis
4 = End stage renal disease on dialysis
5 = Heart failure
6 = Other provider documented reason: breastfeeding during the measurement period
7 = Other provider documented reason: woman of childbearing age not actively taking birth control during the measurement period
8 = Other provider documented reason: allergy to statin
9 = Other provider documented reason: drug interaction (valid drug- drug interactions include HIV protease inhibitors, nefazodone, cyclosporine, gemfibrozil, and danazol)
10 = Other provider documented reason: intolerance (with supporting documentation of trying a statin at least once within the last five years). Additionally, Myopathy and Myositis (CHOL-05) Value Set may be used to document intolerance to statins.

If none of the above contraindications or exceptions are documented, leave BLANK.

NOTE: Items 1 – 5 above can be defined by diagnosis codes that may be used in data collection. Value Sets include: Pregnancy V/Z Codes (PREG-01), Pregnancy Diagnosis Codes (PREG-02), Liver Disease (CHOL-01), Rhabdomyolysis (CHOL-02), ESRD on Dialysis (CHOL-03), and Heart Failure (CHOL-04)

Statin Medication Exception Date:
If the patient has a documented contraindication or exception enter the date of the contraindication or exception. If only the month and year are known, enter the first day of the month.
* Do NOT enter a date that occurred in yyyy. A date in yyyy (beyond measurement period) will create an ERROR upon submission.

ASPIRIN/ANTIPLATELET COMPONENT
Aspirin or Anti-platelet Medication [Numeric] AND
Aspirin or Anti-platelet Date [Date (mm/dd/yyyy)] AND/OR
Aspirin or Anti-platelet Exception [Numeric] AND
Aspirin or Anti-platelet Exception Date [Date (mm/dd/yyyy)]
Numerator component calculation: numerator component compliant if indicated on daily aspirin or anti-platelet medication (prescribed/ ordered) or documented contraindication/exception is present.

Aspirin or Anti-platelet Medication
Enter the code that corresponds to whether the patient is prescribed a daily aspirin product or antiplatelet medication or if an aspirin product or anti-platelet medication was active on the patient's medication list at any time during the measurement period. Please see Appendix D for methods to identify appropriate aspirin products or antiplatelet medications.
1 = Yes, patient was prescribed a daily aspirin product or antiplatelet medication or one was reviewed and active on the patient's medication list.
2 = No, patient was not prescribed a daily aspirin product or antiplatelet medication and one was not reviewed and active on the patient's medication list.
Aspirin/narcotic combination medications do not qualify as a daily aspirin product.
Blank fields will cause an ERROR upon submission.

Aspirin or Anti-platelet Medication Date
Enter the date of the most recent daily aspirin product or anti-platelet medication prescription, order or review of an active medication list that included a daily aspirin product or anti-platelet medication during the measurement period.
* Do NOT enter a date that occurred in yyyy. A date in yyyy (beyond measurement period) will create an ERROR upon submission.
* If a daily aspirin product or anti-platelet medication was not prescribed, ordered or reviewed as an active medication during the measurement period, leave blank.

Aspirin or Anti-platelet Medication Exception
For patients who were not prescribed or taking a daily aspirin product or anti-platelet medication during the measurement period,

enter the code that corresponds to any of the following contraindications or exceptions:
1 = Prescribed anti-coagulant medication during the measurement period
2 = History of gastrointestinal bleeding
3 = History of intracranial bleeding
4 = Bleeding disorder
5 = Other provider documented reason: allergy to aspirin or anti-platelets
6 = Other provider documented reason: use of non-steroidal anti-inflammatory agents
7 = Other provider documented reason: documented risk for drug interaction
8 = Other provider documented reason: uncontrolled hypertension (systolic blood pressure greater than 180 mmHg and/or diastolic blood pressure greater than 110 mmHg)
9 = Other provider documented reason: gastroesophageal reflux disease (GERD)
If none of the above contraindications or exceptions are documented, leave BLANK.
NOTE: Items 1 and 2 above can be defined by diagnosis codes that may be used in data collection. Value Sets include: GI Bleed (ASA-01) and Intracranial Bleed (ASA-02).

Aspirin or Anti-platelet Exception Date
If the patient has a documented contraindication or exception enter the date of the contraindication or exception. If only the month and year are known, enter the first day of the month.
* Do NOT enter a date that occurred in yyyy. A date in yyyy (beyond measurement period) will create an ERROR upon submission.

TOBACCO COMPONENT
Tobacco Status Documentation Date [Date (mm/dd/yyyy)] AND
Tobacco Status [Numeric]
Numerator component calculation: numerator component compliant if tobacco status within the last two years and status is tobacco-free.

Tobacco Status Documentation Date:
Enter the most recent date that the patient's tobacco status was documented during the measurement period or year prior.
* Do NOT enter a date that occurred in yyyy. A date in yyyy (beyond measurement period)will create an ERROR upon submission.
* If the patient's tobacco status is not documented or the date of the documentation cannot be determined, leave BLANK.

Tobacco Status:
Enter the code that corresponds to the patient's most recent tobacco status during the measurement period or year prior.
1 = Tobacco free (patient does not use tobacco; patient was a former user and is not a current user)
2 = No documentation
3 = Current tobacco user (tobacco includes any amount of cigarettes, cigars, pipes or smokeless tobacco)
* If the date of the tobacco status documentation is not documented in the patient record, enter 2.
* E-cigarettes are not considered tobacco products.
A blank field will create an ERROR upon submission.

S.7. **Denominator Statement** *(Brief, narrative description of the target population being measured)*
Patients ages 18 to 75 with ischemic vascular disease who have at least two visits for this diagnosis in the last two years (established patient) with at least one visit in the last 12 months.

S.8. **Target Population Category** *(Check all the populations for which the measure is specified and tested if any):*
Populations at Risk

S.9. **Denominator Details** *(All information required to identify and calculate the target population/denominator such as definitions, specific data collection items/responses , code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
Please also refer to all code lists included in the data dictionary attached in S.2b.
Eligible Specialties:
Family Medicine, Internal Medicine, Geriatric Medicine, Cardiology
Eligible Providers:
Medical Doctor (MD), Doctor of Osteopathy (DO), Physician Assistant (PA), Advanced Practice Registered Nurses (APRN)
Ages:

* 18-75 years of age as of January 1 of the measurement period
Established Patient with Diagnosis:
* Patients are identified as having a diagnosis of ischemic vascular disease (IVD) if they've had at least two face-to-face visits with an eligible provider in an eligible specialty with a diagnosis of IVD (Ischemic Vascular Disease Value Set) during the current or prior measurement period
Event:
* At least one face-to-face visit with an eligible provider in an eligible specialty for any reason during the measurement period

S.10. **Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*
The following exclusions are allowed to be applied to the eligible population: permanent nursing home residents, receiving hospice or palliative care services, died or diagnosis coded in error.

S.11. **Denominator Exclusion Details** *(All information required to identify and calculate exclusions from the denominator such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
* Patient was a permanent nursing home resident at any time during the measurement period
* Patient was in hospice or receiving palliative care at any time during the measurement period
* Patient died prior to the end of the measurement period
* Documentation that diagnosis was coded in error

S.12. **Stratification Details/Variables** *(All information required to stratify the measure results including the stratification variables, definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b)*
The measure for the ischemic vascular disease population is not currently stratified when publicly reported on our consumer website, MN HealthScores. The data is, however, stratified by insurance product in our 2014 Health Care Disparities Report, a hard copy report available on our corporate website at http://mncm.org/wp-content/uploads/2015/03/2014-Health-Care-Disparities-Report-Final.pdf. This report notes a gap in outcomes of ten percentage points between ischemic vascular disease patients in public programs versus other purchasers.

S.13. **Risk Adjustment Type** (Select type. Provide specifications for risk stratification in S.12 and for statistical model in S.14-15)
Statistical risk model
If other:

S.14. **Identify the statistical risk model method and variables** *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development and testing should be addressed with measure testing under Scientific Acceptability)*
The statistical risk model is one of Actual to Expected methodology and is estimated using a logistic model implemented in SAS Procedure Glimmix that accounts for the measure's non-continuous (binary) nature.
Actual to Expected methodology is where the actual measure result remains unaltered, instead a risk adjusted comparison is created based on same proportions of the risk factors that the clinic has.
With Actual to Expected, since the expected is not a stable variable for all clinics, it is not valid to compare the clinic's confidence interval to the expected value. Instead to test whether or not there was a statistically significant difference between the expected value and the actual value achieved by the clinic, a one population proportions test was used. This method is employed to test the proportion of optimally managed patients attributed to a clinic compared to a specified value for that clinic. In the MNCM case the specified value is an expected rate calculated taking into account the overall state rate and adjusted for risk factors specific to the measure.

Variables available to testing in a risk adjustment model for this measure include several demographic variables (age, gender, zip, and insurance product as a proxy for socioeconomic status) and clinical variables (depression and diabetes). Currently, only age and product have the statistical strength (t-test) to be included in the risk adjustment model MNCM is evaluating race/ethnicity, country of origin, primary language as variables in the next year.

(See data dictionary Tab = Risk Adjustment).

S.15. **Detailed risk model specifications** *(must be in attached data dictionary/code list Excel or csv file. Also indicate if available at measure-specific URL identified in S.1.)*
Note: Risk model details (including coefficients, equations, codes with descriptors, definitions), should be provided on a separate

worksheet in the suggested format in the Excel or csv file with data dictionary/code lists at S.2b.
Available in attached Excel or csv file at S.2b

S.15a. **Detailed risk model specifications** (*if not provided in excel or csv file at S.2b*)

S.16. **Type of score:**
Rate/proportion
If other:

S.17. **Interpretation of Score** *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*
Better quality = Higher score

S.18. **Calculation Algorithm/Measure Logic** *(Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.)*
This measure is calculated by submitting a file of individual patient values (e.g. blood pressure, tobacco status, etc) to a HIPAA secure data portal. Programming within the data portal determines if each patient is a numerator case and then a rate is calculated for each clinic site. Please also refer to the measure calculation algorithms submitted within the data dictionary for this measure.
If any component of the numerator is noncompliant for any one of the four components, then the patient is numerator noncompliant for the composite patient level all-or none optimal vascular care measure.

Numerator logic is as follows:

Blood Pressure Component:
Is Blood Pressure date in the measurement year? If no, is numerator noncompliant for this component. If yes, assess next variable.
BP Systolic < 140? If no, is numerator noncompliant for this component. If yes, assess next variable.
BP Diastolic < 90? If yes, is numerator compliant for this component. If no, is numerator noncompliant for this component.
Note: BP needs to occur during the measurement year AND most recent BP systolic less than 140 AND BP diastolic less than 90
Assess next component.

Cholesterol Statin Use Component:
Is the patient on a statin medication? If yes, and most recent date is in the measurement year, is numerator compliant for this component. If no, assess next variable.
For patients not on a statin the following variables are used to assess numerator compliance related to contraindications or exceptions to statin use:
Is the patient age 18 to 20? If yes, numerator compliant (free-pass), if no, assess next variable.
Patients age 21 to 75. Is their most recent LDL in the last five years less than 40? If Yes, numerator compliant (free-pass), if no, assess next variable.
Does the patient have a valid contraindication/ exception to statin use defined as one of the following: pregnancy, active liver disease, rhabdomyolysis, ends stage renal disease on dialysis, heart failure, breastfeeding, allergy to statin, drug-drug interaction with statin, or intolerance with documentation of trying a statin at least once in the last 5 years)? If yes, is numerator compliant for this component. If no, fail this numerator component and remains in the denominator.
Note: Patient is either on a statin (prescribed/ ordered) during the measurement year or has a valid exception either by age, presence or absence of ischemic vascular disease, low untreated LDL or valid contraindication/ exception.
Assess next component.

Tobacco-Free Component:
Is Tobacco Status = 1 (Tobacco Free) and Tobacco Assessment Date a valid date? If yes, is numerator compliant for this component. If no, is numerator noncompliant for this component. Assess next component.

Daily Aspirin/ Anti-platelet Component:
Is the patient on daily aspirin or an antiplatelet? If yes, and date of most recent aspirin/ anti-platelet is in the measurement year is numerator compliant, if no, assess next variable.
Does the patient have a valid contraindication/ exception to aspirin anti-platelet use defined as one of the following: anti-coagulant medication, history of gastrointestinal bleed, history of intracranial bleed, allergy, or physician documented reasons related to: risk

of drug interaction, use of NSAIDS, uncontrolled HTN or gastro-intestinal reflux disease. If yes, is numerator compliant for this component. If no, fail this numerator component and remains in the denominator.
Note: Patients are either on daily aspirin (indicated/ prescribed/ ordered) or an anti-platelet prescribed/ ordered) during the measurement year or has a valid contraindication/ exception.

If all of the above numerator components are in compliance, then the patient calculated as a numerator case for the optimal vascular care measure.

S.19. **Calculation Algorithm/Measure Logic Diagram URL or Attachment** *(You also may provide a diagram of the Calculation Algorithm/Measure Logic described above at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
Available at measure-specific web page URL identified in S.1

S.20. **Sampling** *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*
IF a PRO-PM, identify whether (and how) proxy responses are allowed.
Medical groups are encouraged to submit their full population of patients when possible (EMR) however clinics who are on a paper chart system are allowed to create a random sample of no less than 60 patients per clinic site. MNCM recommends that medical groups submit total population for each measure. By submitting total population, the confidence interval around the rate narrows, indicating a higher confidence that the rate accurately reflects the clinics' performance. If total population is not an option for a medical group, MNCM encourages medical groups to submit a large sample. The minimum required sample is 60 patients per clinic site, per measure. If a clinic site has less than 60 patients in the total population for the measure, the entire population must be submitted.

For 2015 dates of service 99.5% submitted total population, 0.5% submitted a sample.

Excel's Random Number Generator Instructions:
For lists generated in Excel, use the "RAND" function to assign a random number to each record (please also see Microsoft Excel Help, topic RAND for more information):

1. Insert a blank column on the leftmost side of the spreadsheet
2. Label new column "RAND"
3. Place cursor in the first blank cell (A2) and type =RAND()
4. Press enter (a number like 0.793958 will appear)
5. Place the cursor back into this cell; resting over the corner to have the pointer change to a black cross, double click or drag the formula down to the last row/patient
6. Highlight the whole column and click Edit, Copy, Paste Special = Values to freeze the random number (otherwise it will change with every click on the spreadsheet)
7. Sort entire patient population by this new random number
8. Work down the list row by row, starting with row 1 until the number of records in the sample is met for submission (at least 60 patients per clinic, per measure)
9. If a patient meets one of the accepted exclusions, keep working down the list and use oversamples that are after the number of records in the sample. For example, if 60 records will be submitted and 2 exclusions were found, include patient rows 61 and 62 to replace the excluded records.

S.21. **Survey/Patient-reported data** *(If measure is based on a survey, provide instructions for conducting the survey and guidance on minimum response rate.)*
IF a PRO-PM, specify calculation of response rates to be reported with performance measure results.

S.22. **Missing data** (specify how missing data are handled, e.g., imputation, delete case.)
Required for Composites and PRO-PMs.
For this patient level all-or-none composite measure, elements missing from any component (e.g. visit but no blood pressure during the measurement year) are counted as a numerator component fail and therefore the patient would be accounted for and remain in the denominator.

S.23. **Data Source** *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*
If other, please describe in S.24.

Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

S.24. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument e.g. name of database, clinical registry, collection instrument, etc.)*
IF a PRO-PM, identify the specific PROM(s); and standard methods, modes, and languages of administration.
An excel template with formatted columns for data fields is provided. Many medical groups extract the information from their EMR. Registries can be used as a source of information to create the data file; however groups must ensure that all of their eligible patients are included. Paper abstraction forms are provided for those clinics who wish to use them as an interim step to creating their data file. All data is uploaded in electronic format (.csv file) to a HIPAA secure, encrypted and password protected data portal.

S.25. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
Available at measure-specific web page URL identified in S.1

S.26. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*
Clinician : Group/Practice

S.27. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*
Ambulatory Care : Clinician Office/Clinic
If other:

S.28. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*
Calculation of the patient level all-or-none composite measure is indicated 1) in the measure algorithms provided in the data dictionary. Please note that while the all-or-none composite measure is considered to be the gold standard, reflecting best patient outcomes, the individual components may be measured as well. This is particularly helpful in quality improvement efforts to better understand where opportunities exist in moving the patients toward achieving all of the desired outcomes. Individual component logic is included below:

Denominator is the same for calculating individual component rates as the patient level all-or-none composite measure: Patients ages 18 to 75 with ischemic vascular disease who have at least two visits for this diagnosis in the last two years (established patient) with at least one visit in the last 12 months. Exclusions are: permanent nursing home resident, hospice or palliative care, death, and documentation diagnosis coded in error.

Component for Blood Pressure Control:
Is the BP date in the measurement year? If No, fails the numerator. If Yes, assess next variable.
Is the most recent BP value less than 140 systolic AND less than 90 diastolic? If Yes, is in the numerator for this component.
Expressed as a rate:
Patients with most recent BP during the measurement year is less than 140 systolic AND 90 diastolic/
Eligible patients with ischemic vascular disease

Component for Cholesterol/ Statin Use:
Is the patient on a statin medication? If yes, and most recent date is in the measurement year, is in the numerator for this component.
For patients not on a statin the following variables are used to assess numerator compliance related to contraindications or exceptions to statin use:
Is the patient age 18 to 20? If yes, in the numerator (free-pass), if no, assess next variable.
Patient age 21 to 75- Is their most recent LDL in the last five years less than 40? If Yes, in the numerator (free-pass), if no, assess next variable.
Does the patient have a valid contraindication/ exception to statin use defined as one of the following: pregnancy, active liver disease, rhabdomyolysis, ends stage renal disease on dialysis, heart failure, breastfeeding, allergy to statin, drug-drug interaction with statin, or intolerance with documentation of trying a statin at least once in the last 5 years)? If yes, is in the numerator. If no, fail this numerator component and remains in the denominator.
Expressed as a rate:
Patients with statin use unless with contraindications/ exceptions/
Eligible patients with ischemic vascular disease

Component for Tobacco-Free:
Is the date of smoking status in the measurement year or the year prior? If No, fails the numerator. If Yes, assess next variable.
Is the patient's tobacco status noted as tobacco-free? If Yes, is in the numerator.
Expressed as a rate:
Patients with most recent tobacco status during the measurement year or the year prior is free of all tobacco products (tobacco free)/ Eligible patients with ischemic vascular disease

Component for Daily Aspirin/ Anti-platelet Component:
Is the patient on daily aspirin or an antiplatelet? If yes, and date of most recent aspirin/ anti-platelet is in the measurement year is numerator compliant, if no, assess next variable.
Does the patient have a valid contraindication/ exception to aspirin anti-platelet use defined as one of the following: anti-coagulant medication, history of gastrointestinal bleed, history of intracranial bleed, allergy, or physician documented reasons related to: risk of drug interaction, use of NSAIDS, uncontrolled HTN or gastro-intestinal reflux disease. If yes, is numerator compliant for this component. If no, fail this numerator component and remains in the denominator.
Expressed as a rate:
Patients with daily aspirin/ anti-platelet use unless with contraindications/ exceptions/Eligible patients with ischemic vascular disease

2a. Reliability – See attached Measure Testing Submission Form
2b. Validity – See attached Measure Testing Submission Form
0076_OVC_Template_MeasSubm_CompositeMeasTesting_MNCM_2016-635959802775191129.docx

NATIONAL QUALITY FORUM—Composite Measure Testing (subcriteria 2a2, 2b2-2b7, 2d)

Measure Number (*if previously endorsed*)**: 0076**
Composite Measure Title: Optimal Vascular Care
Date of Submission: 4/11/2016
Composite Construction:
☐Two or more individual performance measure scores combined into one score
☒ All-or-none measures (e.g., all essential care processes received or outcomes experienced by each patient)
☐ Any-or-none measures (e.g., any or none of a list of adverse outcomes experienced, or inappropriate or unnecessary care processes received, by each patient)

Instructions: Please contact NQF staff before you begin.

- If a component measure is submitted as an individual performance measure, the non-composite measure testing form must also be completed and attached to the individual measure submission.
- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- **For all composite measures, sections 1, 2a2, 2b2, 2b3, 2b5, and 2d must be completed.**
- **For composites with outcome and resource use measures**, section **2b4** also must be completed.
- If specified for **multiple data sources/sets of specificaitions** (e.g., claims and EHRs), section **2b6** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2), validity (2b2-2b6), and composites (2d) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 25 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **PRO-PMs and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b2. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **PRO-PMs and composite performance measures**, validity should be demonstrated for the computed performance score.

2b3. Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; [12]

AND

If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, 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). [13]

2b4. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and sociodemographic factors) that influence the measured outcome and are present at start of care; [14] and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b5. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** [15] **differences in performance**;

OR

there is evidence of overall less-than-optimal performance.

2b6. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b7. For **eMeasures, composites, and PRO-PMs** (or other measures susceptible to missing data), analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

2d. For composite performance measures, empirical analyses support the composite construction approach and demonstrate that:

2d1. the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and

2d2.the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.

(*if not conducted or results not adequate, justification must be submitted and accepted)*

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements 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 of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality.

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

13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions.
15. 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 percent v. 75 percent) 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 less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE
Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation. If different data sources are used for different components in the composite, indicate the component after the checkbox.*)

Measure Specified to Use Data From: (***must be consistent with data sources entered in S.23***)	**Measure Tested with Data From:**
☒ abstracted from paper record	☒ abstracted from paper record
☐ administrative claims	☐ administrative claims
☐ clinical database/registry	☐ clinical database/registry
☒ abstracted from electronic health record	☒ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).
In 2010 (2009 dates of service), 128 medical groups representing 573 physician clinics and 95,791 patients with IVD in Minnesota and neighboring communities submitted data for this measure. Of the 95,791 IVD patients, a sample of 63,241 patients was submitted for rate calculation. 79% of the clinics submitted full population data, 21% of clinics submitted a random sample. Dates of service included 01/01/2009 to 12/31/2009 (LDL date of service was a 15-month time frame 10/01/2008 to 12/31/2009).

The data submitted represents 66% of all eligible patients; based on the large sample size, the results can be reliably reproduced. The data submission process requires individual patient data for each component of the "all or none" composite measure (e.g., most recent LDL value and blood pressure in the measurement period). This information is accurately captured as evidenced by post submission validation audits against the patient's medical record.

[2016 with 2015 Dates of Service]
Existing data is used. Data is collected and reported on an annual basis for this measure in MN and surrounding border communities. In 2016 (2015 dates of service), 111 medical groups representing 671 physician clinics and 104,494 patients with IVD in Minnesota and neighboring communities submitted data for this measure. Of the 104,494 IVD patients, 104,395 patients were submitted for rate calculation. 99.6% of the clinics submitted full population data, with only 3 clinics submitting a random sample. Dates of service included 01/01/2015 to 12/31/2015.

The data submitted represents 99.9% of all eligible patients; based on the large sample size, the results can be reliably reproduced. The data submission process requires individual patient data for each component of the "all or none" composite measure (e.g., most recent blood pressure values in the measurement period). This information is accurately captured as evidenced by post submission validation audits against the patient's medical record.
Types of fields included in the submission for 2015 dates of service include the following:

Clinic ID ▪ Patient ID ▪ Patient Date of Birth ▪ Patient Gender Zip Code, Primary Residence ▪ Race/Ethnicity1 ▪Race/Ethnicity2 ▪ Race/Ethnicity3 ▪ Race/Ethnicity4 ▪ Race/Ethnicity5 ▪ Country of Origin Code ▪ Country of Origin "Other" Description ▪ Preferred Language Code ▪ Preferred Language "Other" Description ▪ Provider NPI ▪ Provider Specialty Code ▪ Insurance Coverage Code ▪ Insurance Coverage "Other" Description ▪Insurance Plan Member ID ▪Patient Has Diabetes? ▪ Patient Has Depression? ▪ LDL Date ▪ LDL Value ▪ BP Date ▪ BP Systolic ▪ BP Diastolic ▪ Statin Medication ▪ Statin Medication Date ▪ Statin Medication Exception ▪ Statin Medication Exception Date▪ Aspirin or Anti-platelet Medication ▪ Aspirin or Anti-platelet Medication Date ▪ Aspirin or Anti-platelet Medication Exception ▪ Aspirin or Anti-platelet Medication Exception Date ▪ Tobacco Status Documentation Date ▪ Tobacco Status

1.3. What are the dates of the data used in testing? 1/1/2009 to 12/31/2009 and current year 1/1/2015 to 12/31/2015

1.4. What levels of analysis were tested? (*testing must be provided for <u>all</u> the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: **(*must be consistent with levels entered in item S.26*)**	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☒ group/practice	☒ group/practice
☐ hospital/facility/agency	☐ hospital/facility/agency
☐ health plan	☐ health plan
☐ other: Click here to describe	☐ other: Click here to describe

1.5. How many and which <u>measured entities</u> were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)
[2010 with 2009 Dates of Service]
Characteristics of the entities reporting data:
Based on number of physicians, the size of the 128 medical groups that submitted data ranged from one-physician practices to medical groups with more than 2,700 physicians. Ranges include: Medical groups with <25 physicians = 87; medical groups with 25-99 physicians = 25; medical groups with 100-249 physicians = 5; medical groups with 250+ physicians = 11. 50 medical groups were located within the Twin Cities metro area, while 78 medical groups were located outside of the Twin Cities metro area. 110 medical groups were identified as primary care clinics, 17 medical groups were identified as multi-specialty clinics, and one group was identified as a single-specialty clinic (cardiology).

Of the 573 clinic sites that reported data, 455 clinics used an electronic medical record in some capacity for the clinical data collection (data extraction/query, or manual data abstraction), and 118 clinics used paper records for the clinical data collection.

[2016 with 2015 Dates of Service]
Characteristics of the entities reporting data:
Based on number of physicians, the size of the 111 medical groups that submitted data ranged from one-physician practices to medical groups with more than 500 physicians. Ranges include: Medical groups with <25 physicians = 80; medical groups with 25-99 physicians = 15; medical groups with 100-249 physicians = 9; medical groups with 250+ physicians = 7. 39 medical groups were located within the Twin Cities metro area, while 72 medical groups were located outside of the Twin Cities metro area. 52 medical groups were identified as primary care clinics, 52 medical groups were identified as multi-specialty clinics with 22 inclusive of cardiology, and one group was identified as a single-specialty clinic (cardiology).

Of the 671 clinic sites that reported data, 57% used their EMR exclusively to extract data, 40% supplemented data extraction with some chart abstraction and 3% relied on manual abstraction processes of their EMR or paper based record. Interestingly, in 2010 118 clinics in MN had paper based system, in 2015 only 5 clinics are on a paper based system.

1.6. How many and which <u>patients</u> were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)
In 2010 (2009 dates of service), 128 medical groups representing 573 physician clinics and 95,791 patients with IVD in Minnesota and neighboring communities submitted data for this measure. Of the 95,791 IVD patients, a sample of 63,241 patients was submitted for rate calculation.

[2016 with 2015 Dates of Service]
111 medical groups representing 671 physician clinics and 104,494 patients with IVD in Minnesota and neighboring communities submitted data for this measure. Of the 104,494 IVD patients, 104,395 patients were submitted for rate calculation. 99.6% of the clinics submitted full population data, with only 3 clinics submitting a random sample.

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.
There are no differences.

1.8 What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example: patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).
Insurance product as a proxy for socioeconomic status. Insurance Coverage Code is included in the patient level file that is submitted from the medical group and is translated to a higher level insurance product (commercial, Medicare, MHCP- state public program and uninsured). Insurance product has demonstrated properties for inclusion in a risk adjustment model (t-tests < 0.01 to 0.02).

2a2. RELIABILITY TESTING

2a2.1. What level of reliability testing was conducted?
<u>Note</u>: *Current guidance for composite measure evaluation states that reliability must be demonstrated for the composite performance measure score.*
☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. Describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)
Used paper "Reliability in Provider Profiling" by John L. Adams, Ph.D as a reference
The BETABIN macro was used on each measure (SAS).

- First, we need to find the provider-to-provider variance:
  - $\sigma^2 = (\alpha\ \beta) / (\alpha + \beta + 1)(\alpha + \beta)^2$
  - $= (8.195 * 10.324) / (8.195 + 10.324 + 1)(8.195 + 10.324)^2$
  - = **0.0126** (plug this value into the reliability equation)
- Reliability = $\sigma^2 / (\sigma^2 + (p(1 - p)/n))$
  - p = rate
  - n = number of eligible patients
- Determine reliability rate for each provider.
- Average the reliability rate.

2a2.3. What were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)
Reliability = 0.9068

BETABIN Macro: Simple Binomial Model

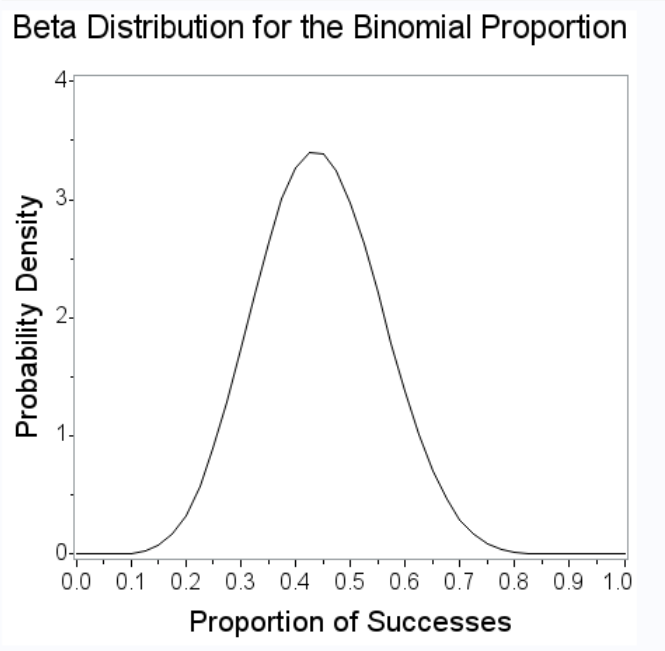
parameter	Estimate	Standard Error	t Value	Prob > |t|	Alpha	Lower	Upper
mu	0.4664	0.0010	473.12	<.0001	0.05	0.4644	0.4683
|mu-0.5|	0.0336	0.0010	34.13	<.0001	0.05	0.0317	0.0356

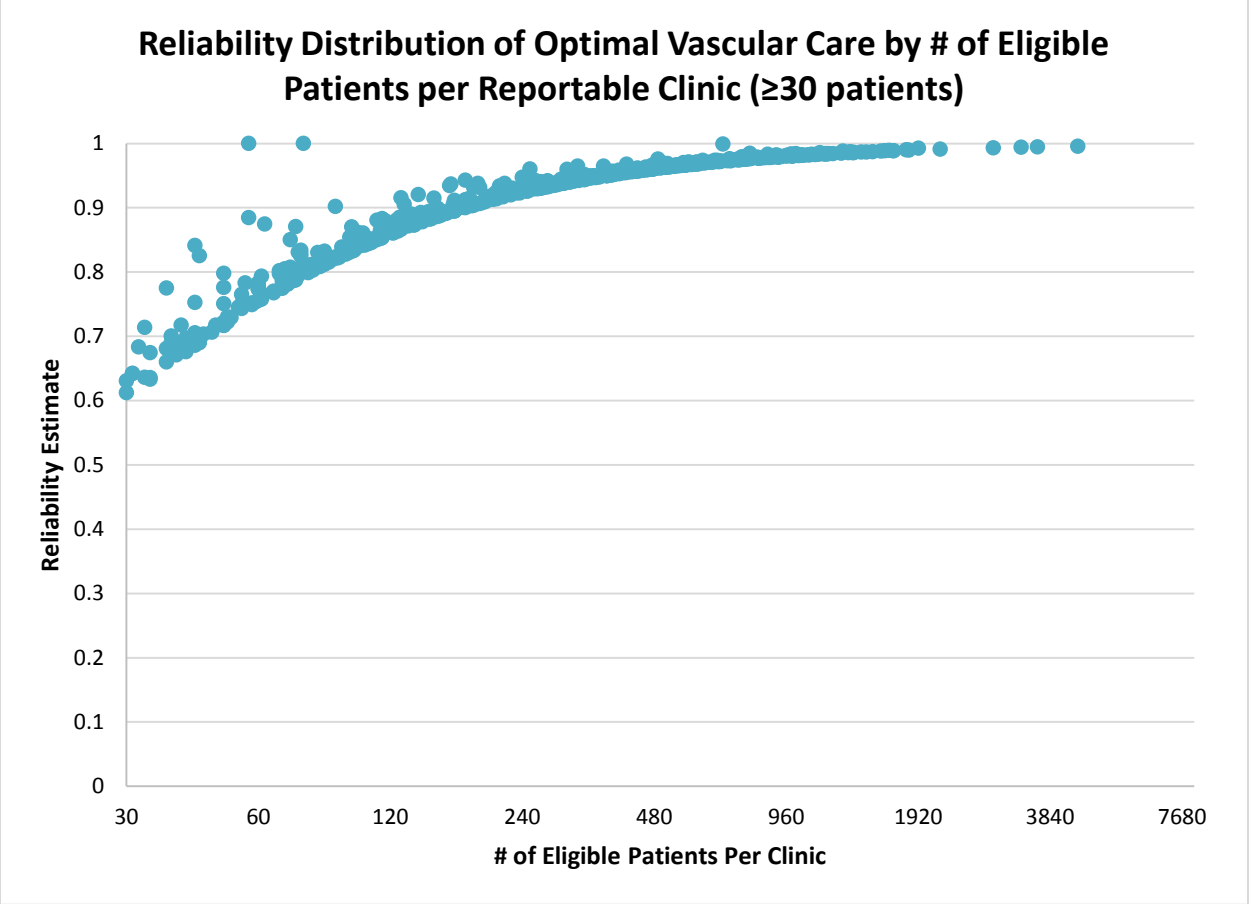
BETABIN Macro: Beta-Binomial Model Parameters

parameter	Estimate	Standard Error	t Value	Pr > |t|	Alpha	Lower	Upper
mu	0.4425	0.004777	92.65	<.0001	0.05	0.4332	0.4519
alpha	8.1950	0.5307	15.44	<.0001	0.05	7.1548	9.2351
beta	10.3237	0.6644	15.54	<.0001	0.05	9.0215	11.6260
gamma	0.05123	0.003102	16.52	<.0001	0.05	0.04515	0.05731
theta	0.05400	0.003446	15.67	<.0001	0.05	0.04725	0.06075
|mu-0.5|	0.05748	0.004777	12.03	<.0001	0.05	0.04811	0.06684

BETABIN Macro: Variance-Covariance Matrix of Estimated Parameters

Label	mu	alpha	beta	gamma	theta
mu	0.000023	0.000431	-0.00041	-5.16E-8	-5.73E-8
alpha	0.000431	0.2816	0.3367	-0.00162	-0.00180
beta	-0.00041	0.3367	0.4414	-0.00204	-0.00227
gamma	-5.16E-8	-0.00162	-0.00204	9.622E-6	0.000011
theta	-5.73E-8	-0.00180	-0.00227	0.000011	0.000012





2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

In terms of understanding reliability in detecting signal to noise, a reliability score of 0.70 or greater is considered acceptable for drawing conclusions about groups. This data analysis, along with precise specifications and excellent validation results of critical data elements, demonstrates this measure construct to be reliable and detect meaningful differences among provider groups.

2b2. VALIDITY TESTING

Note*: Current guidance for composite measure evaluation states that validity should be demonstrated for the composite performance measure score. If not feasible for initial endorsement, acceptable alternatives include assessment of content or face validity of the composite OR demonstration of validity for each component. Empirical validity testing of the composite measure score is expected by the time of endorsement maintenance.*

2b2.1. What level of validity testing was conducted?

☒ **Composite performance measure score**

- ☒ **Empirical validity testing**
- ☐ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)
- ☐ **Systematic assessment of content validity**

☒ **Validity testing for component measures** *(check all that apply)*

Note*: applies to ALL component measures, unless already endorsed or are being submitted for individual endorsement.*

- ☐ **Endorsed (or submitted) as individual performance measures**
- ☒ **Critical data elements** (*data element validity must address ALL critical data elements*)
- ☒ **Empirical validity testing of the component measure score(s)**

☐ **Systematic assessment of face validity of component measure score(s) as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)

2b2.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

Critical Data Element:
[2010 Validation Audit Results:]
Of the 128 medical groups submitting data in 2010, 17 groups initially failed the audit and remedy plans were developed. All 17 groups resubmitted and passed subsequent audit.
Types of Errors Found in Validation Audits: BP was not most recent, EMR did not pull the correct date or value, ASA date could not be validated, ASA date not reported, LDL date not reported or more recent date found, and Tobacco status was not correct.

Validating the submitted data via the direct data submission process is completed in four steps: denominator certification, data quality checks, validation audit, and the two-week medical group review period.

Denominator certification prior to data collection and extraction/ abstraction ensures that all medical groups apply the denominator criteria correctly and in a consistent manner. MNCM staff review the documentation to verify all criteria were applied correctly, prior to approval for data submission.
Denominator certification documentation for this measure includes:

- Date of Birth (ranges)
- Date of Service (ranges)
- ICD-9 Codes used
- Eligible specialties and provider types
- Exclusions to the measure and attest to mechanism for exclusions
- Attestations related to changes in medical record or billing systems
- Supplying all query code for review

Common areas of correction in denominator for this measure included missing query code, incorrect date of birth ranges, incorrect dates for counting visits, missing ICD-9 or ICD-10 codes or incomplete attestation. All were corrected prior to data submission.

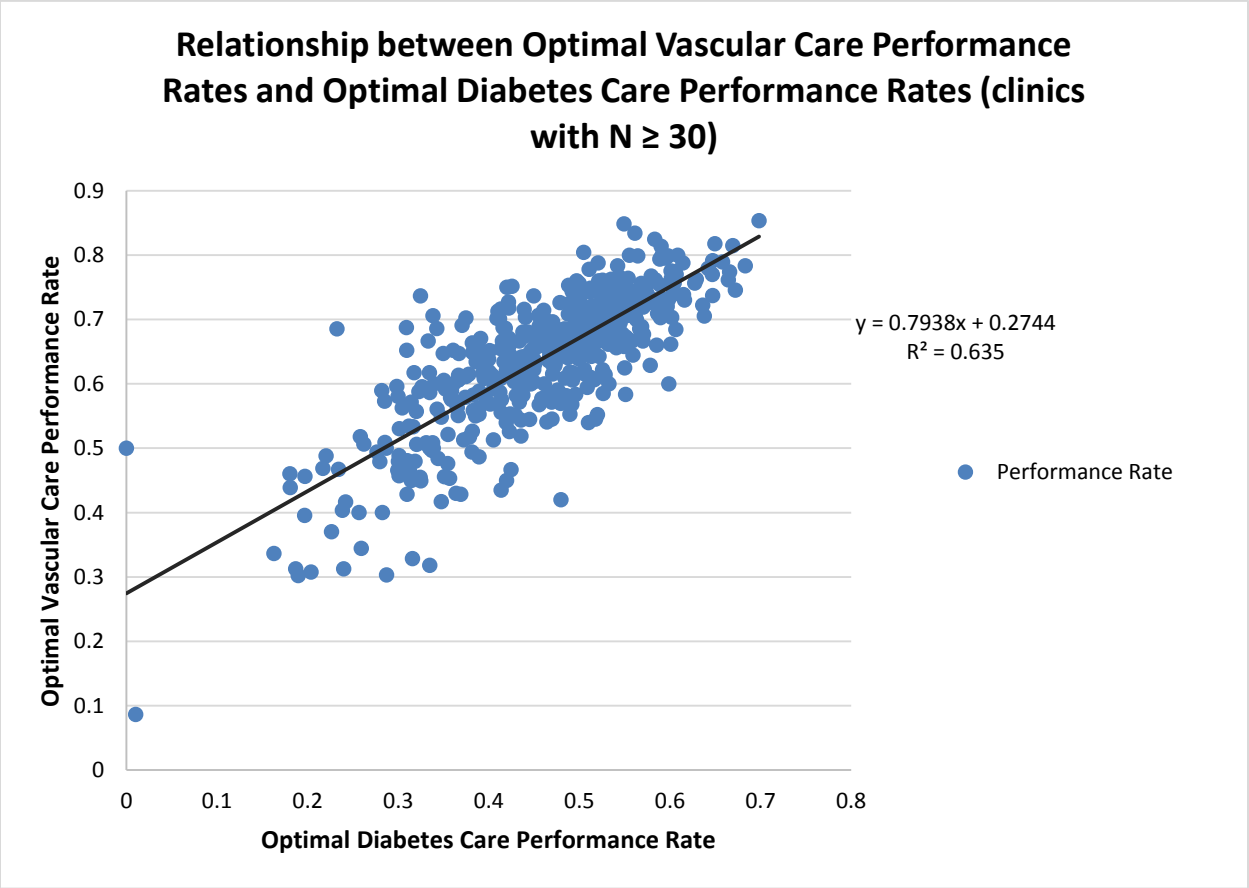
Following data submission to the MNCM Data Portal, there are additional data quality checks in place for evaluating the accuracy of data submitted. During file upload, program checks for valid dates, codes and values and presents users with errors and warnings. Additionally, MNCM staff review population counts (denominator) and outcome rates for any significant variance from the previous year's submission and may prompt further clarification from the medical group.

Validation audits verify that the clinical data submitted for the numerator component of the measure matched the data in the patient record. Other data elements are also audited to verify the patient was included in the denominator correctly (e.g., diagnosis of ischemic vascular disease).

Critical Data Element
[2015 Validation Audit Results:]
In 2015, for the vascular measure, MNCM audited 123 medical groups; 80% of those submitting data. 81% passed the initial audit, 19% required a correction plan and all re-submitted their data and passed the audit with ≥ 90% accuracy. Types of discrepancies noted on audit included: aspirin date during the measurement period, tobacco status, date of birth errors, most recent blood pressure, and inclusion of patient without the diagnosis of IVD.

Validity was tested for the computed composite score by testing the correlation of medical group performance with their performance on the Optimal Diabetes Care measure (NQF# 0729). Ischemic vascular disease and diabetes are chronic conditions that require ongoing management of multiple risk factors in order to reduce a patient's overall risk of developing long term complications. It is expected that the quality of care provided by a medical group to patient with ischemic vascular disease would be of similar quality as the care provided to patients with diabetes, and the respective performance measure scores should demonstrate such.

2b2.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)
Based on linear regression analysis, a medical group's performance on the Optimal Vascular Care measure is associated with its performance on the Optimal Diabetes Care measure, as demonstrated by an r^2 value of 64%, representing a fairly strong correlation.



2b2.4. What is your interpretation of the results in terms of demonstrating validity? (i.*e., what do the results mean and what are the norms for the test conducted?*)
High compliance with critical data element validity as demonstrated by annual validation audit processes.

As demonstrated by the r^2 value, 64% of the total variation in performance on the Optimal Vascular Care measure can be explained by variation in the Optimal Diabetes Care measure. This degree of correlation indicates that the Optimal Vascular Care composite measure score accurately reflects the quality of care provided.

2b3. EXCLUSIONS ANALYSIS
Note: *Applies to the composite performance measure, as well all component measures unless they are already endorsed or are being submitted for individual endorsement.*
NA ☐ no exclusions — *skip to section 2b4*

2b3.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)
[2010 Exclusion Testing Results:]
In addition to the denominator certification process that describes how groups excluded patients, we asked groups to record all the individual patients that they excluded and the reasons for the exclusions. Groups submitted a list of excluded patients to MNCM. The total number of exclusions submitted (n = 1,403) in 2010 was 2.2% of the number of patients submitted (1,403/63,241). Clinics that

submitted excluded patients most often manually documented exclusions upon record review. Some clinics with an EMR were also able to submit patients that they were able to filter out of the patient population (e.g., deceased patients).

If a clinic elected to take allowable exclusions, they were required to submit a list of excluded patients along with the type of exclusion per patient. MNCM conducted a review of all exclusions taken to validate that only allowable exclusions were taken and to identify the number of exclusions by type.

2b3.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)
The frequency of the use of the exclusions under study was 2.2% of the number of patients submitted (1,403/63,241). Medical group utilization of exclusions: 77 of 128 (60%) of groups submitted exclusions.

[2016 Exclusion Results:]
Total number of exclusions: 1,249/104,395 = 1.20%

- 234 nursing home
- 91 hospice
- 918 deceased
- 6 coded in error

Number of medical groups that submitted exclusions: 51 of 113 (45.1%)

2b3.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)
Although exclusions are somewhat rare (1.2% of the population), they are necessary for this measure. The upper age limit cut-off does limit the frail elderly population in which the targets may not be appropriate, but allowing these exclusions serves its purpose to capture potentially frail patients who are less than age 75.

2b4. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
Note: *Applies to all outcome or resource use component measures, unless already endorsed or are being submitted for individual endorsement.*
If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b4.1. What method of controlling for differences in case mix is used? *(check all that apply)*
☐ **Endorsed (or submitted) as individual performance measures**
☐ **No risk adjustment or stratification**
☒ **Statistical risk model**
☐ **Stratification by risk categories**
☐ **Other,** Click here to enter description

2b4.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*)
The selection of risk factors for adjustment of clinical quality performance measure scores is directed by a framework of criteria that must be considered for each factor[i].
The following criteria are to be applied during:

1.) Measure development when recommending variables for data collection and testing for potential risk adjustment, and
2.) Selection of tested variables for the application of risk adjustment.
3.) Reevaluation of currently applied risk adjustment factors.

Criteria	Rationale
1. Clinical/conceptual relationship with the outcome of interest	A logical theory must explain the association between the factor and the outcome. Begin with conceptual model informed by research and experience; does not require a direct causal relationship
2. Empirical association with the outcome of interest	A statistical association to confirm the conceptual relationship
3. Variation in prevalence of the factor across the measured entities	If there is no variation in prevalence across providers being measured, it will not bias performance results
4. Not confounded with quality of care – risk factors should: • be present at the start of care and • not represent the quality of care provided (e.g., treatments, interventions, expertise of staff)	Trying to isolate effects of the provider – quality of care Ensures not a result of care provided Although these could explain variation in outcome, trying to isolate differences in performance due to differences in the care provided
5. Resistant to manipulation or gaming – generally, a diagnosis or assessment data (e.g., functional status score) is considered less susceptible to manipulation than a clinical procedure or treatment (e.g., physical therapy)	Ensures validity of performance score as representing quality of care (vs. for example, up coding)
6. Accurate data that can be reliably and feasibly captured at a reasonable cost	Data and resource limitations often represent a practical constraint to what factors are included in risk models.
7. Contribution of unique variation in the outcome (i.e., not redundant or highly correlated with another risk factor)	Prevent over-fitting and unstable estimates, or coefficients that appear to be in the wrong direction, reduce data collection burden
Potentially, improvement of the risk model (e.g., risk model metrics of discrimination – i.e., sensitivity/specificity, calibration) and sustained with cross validation	Change in R-squared or C-statistic may not be significant, but calibration at different deciles of risk might improve May not appear to be a big change but could represent meaningful differences in terms of the outcome (e.g., lives, dollars) Order of entry into a model may influence this result
Potentially, face validity and acceptability	Some factors may not be indicated empirically, but could improve acceptability – need to weigh against negative impact on model, feasibility and burden of data collection

2b4.4a. What were the statistical results of the analyses used to select risk factors?

Initial analysis was conducted on 2012 measure year:

TABLE 1: EFFECT OF POTENTIAL RISK ADJUSTERS ON OVC

Variable	Contrast	Estimate	T-value	Odds Ratio	Lower 95% CI	Upper 95% CI
1A: MODEL WITHOUT SES AND RACE FROM ZIP CODE DATA						
Age						
18-25	66-75	-1.12**	-2.71	0.32**	0.14	0.73
26-50	66-75	-0.63**	-18.68	0.53**	0.50	0.57
51-65	66-75	-0.26**	-12.91	0.77**	0.74	0.80
Gender						
Female	Male	-0.32**	-17.95	0.73**	0.70	0.75
Comorbidity						
Depressed	Not Depressed	-0.09**	-4.37	0.91**	0.87	0.95
Distance from Clinic						
<5 miles	Same Zip	0.02	0.71	1.02	0.97	1.07
5-10 miles	Same Zip	0.00	0.19	1.00	0.96	1.06
10-20 miles	Same Zip	0.02	0.66	1.02	0.97	1.07
20+ miles	Same Zip	-0.13**	-4.74	0.87**	0.83	0.92

Insurance						
Medicare	Commercial	-0.06**	-2.79	0.94**	0.90	0.98
Medicaid / MSHO / Special Needs / Self-pay / Uninsured	Commercial	-0.75**	-28.61	0.47**	0.45	0.50
Constant		-0. 29	-1.01			

2b4.4b. Describe the analyses and interpretation resulting in the decision to select SDS factors (e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects)

While Age, Insurance Product, Gender and Depression all made the initial statistical requirement of significant variation in results, Gender did not show enough variation between clinics to contribute to the unique variation of clinic level results (Criteria #3) and depression was not selected due to relative high cost of collection.(Criteria #7)

Race, Ethnicity, Language and Country of Origin (RELO) did not have a high completion rate across all clinics to be considered for risk adjustment at this time, we are continuing to work with the medical community to achieve the goal of evaluating RELO at the clinic level.

Therefore the risk variables selected were Age and Insurance Product.

2b4.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)
Because this measure is a binary variable (0 or 1), the risk adjustment model was estimated using a logistic model implemented in the SAS Procedure Glimmix that accounts for its non-continuous nature. The risk adjusters and an indicator for each clinic were included in the model. The estimated coefficient for the clinic indicator measures the clinic's ODC adjusting for the patient risk adjusters that were included in the model. The clinic level indicator was used to construct a risk adjusted ODC score at the clinic level that ranged from 0 to 1 (0% to 100%). The effect of risk adjustment on clinic rankings was calculated by comparing the risk adjusted OVC to the unadjusted OVC measure, the average ODC for all patients reported by the clinic.

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
if stratified, skip to 2b4.9

2b4.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**
Analysis of Maximum Likelihood Estimates [2014 dates of service]

		n	Optimal Rate	Comparison	t test
Product	Commercial	35,706	71%	Medicare	0.02
	Medicare	55,993	71%		
	MHCP	9,220	54%	Medicare	<.01
	Uninsured	1,735	55%	Medicare	<.01
Age	18- 25	44	68%	66-75	0.29
	26- 50	7,022	58%	66-75	<.01
	51-65	44,105	66%	66-75	<.01
	66 - 75	51,483	74%		

2b4.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):
Test of selected variables

Analysis of Maximum Likelihood Estimates						
Parameter	Comparisoin	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept		1	-0.8067	0.1727	21.8112	**<.0001**
Medicare	Commercial	1	0.3539	0.02	313.1994	**<.0001**
MHCP	Commercial	1	0.7962	0.025	1010.692	**<.0001**
Uninsured	Commercial	1	-0.367	0.0265	192.1117	**<.0001**
18-25	51-65	1	0.1782	0.1703	1.0954	**0.2953**
26 - 50	51-65	1	-0.1394	0.0135	106.0312	**<.0001**
66 - 75	51-65	1	0.2527	0.00939	723.3905	**<.0001**

All results of both variables are significant except for ages less than 26 due to small sample size.

Test of correlation between variables, R square test

Estimated Correlation Matrix							
Parameter	Intercept	Medicare	MHCP	Uninsured	18-25	26 - 50	66 - 75
Intercept	1	(0.046)	(0.050)	(0.128)	(0.986)	(0.074)	(0.064)
Medicare	-0.0458	1.000	0.336	(0.150)	0.002	(0.042)	**0.619**
MHCP	-0.0497	0.336	1.000	(0.115)	0.026	0.072	0.084
Uninsured	-0.1279	(0.150)	(0.115)	1.000	(0.006)	(0.035)	(0.029)
18-25	-0.9855	0.002	0.026	(0.006)	1.000	0.014	0.014
26 - 50	-0.0739	(0.042)	0.072	(0.035)	0.014	1.000	0.169
66 - 75	-0.0637	**0.619**	0.084	(0.029)	0.014	0.169	1.000

R-Square	0.0223	**Max-rescaled R-Square**	0.0315

The only two results that are correlated is being over 65 and being on Medicare, which makes logical sense.

2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

Impact on Clinic Level Measurement Clinic Count				
		With Risk Adjustment		
		Below	Expected	Above	Total
No Risk Adj	Below	65	30	0	84

	Average	0	297	1	298
	Above	0	25	59	95
		65	352	60	477
	Direction of Impact		2015 Impact		*Previous Year*
	Move to Expected	55	11.5%		*13.5%*
	Moved away from Expected	1	0.2%		*0.2%*
	Improved	31	6.5%		*5.4%*
	Worse	25	5.2%		*8.3%*
	Impacted	56	11.7%		*13.8%*

2b4.9. Results of Risk Stratification Analysis:
Risk Segmentation Analysis
Optimal Vascular Care 1/1/2014 – 12/31/2014

Patients					
Age of patient					
	18-25	26 - 50	51 - 65	66 - 75	Total
Insureance Type					
Commercial	15	3,905	25,599	3,771	33,290
Medicare	4	849	9,490	43,154	53,497
State Public Programs	19	1,574	5,255	1,710	8,558
Uninsured	2	290	986	261	1,539
	40	6,618	41,330	48,896	96,884

Optimally Manage Patients					
Age of patient					
	18-25	26 - 50	51 - 65	66 - 75	Total
Insureance Type					
Commercial	12	2,594	18,181	2,828	23,615
Medicare	2	401	5,772	32,074	38,249
State Public Programs	12	688	2,776	1,081	4,557
Uninsured	1	135	533	169	838
	27	3,818	27,262	36,152	67,259

	18-25	26 - 50	51 - 65	66 - 75	Total
Commercial	80.0%	66.4%	71.0%	75.0%	70.9%
Medicare	50.0%	47.2%	60.8%	74.3%	71.5%
State Public Programs	63.2%	43.7%	52.8%	63.2%	53.2%
Uninsured	50.0%	46.6%	54.1%	64.8%	54.5%
	67.5%	57.7%	66.0%	73.9%	69.4%

t test compared to Medicare patient over 65 years old					
	18-25	26 - 50	51 - 65	66 - 75	Total
Commercial	0.44	<.01	<.01	0.26	0.20
Medicare	0.20	<.01	<.01	0.64	0.20
State Public Programs	0.20	<.01	<.01	<.01	0.32
Uninsured	0.32	<.01	<.01	<.01	<.01

Conclusion: For patients under the age of 26, the sample size is too small to make any type of determination. For all ages over 25 and for all insurance products, there are significant differences between the categories.

2b4.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i*.e., what do the results mean and what are the norms for the test conducted?*)
Age and Insurance Product meet all requirements for Risk Variables; clinical concept, empirical association, variation across entries, not confounded with quality of care, resistant to manipulation, feasible and reliable collection and contribute to unique variation therefore are appropriate for risk adjustment.

***2b4.11. Optional Additional Testing for Risk Adjustment** (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE
Note: *Applies to the composite performance measure.*

2b5.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

Methodology:

Identifying High Performing Medical Groups/Clinics

For each measure, both individual medical group rates and a medical group average rate were calculated. Medical groups that achieved high performance were identified by comparing the individual medical group/clinic rate with the statewide average. Medical groups that had rates and 95 percent confidence intervals that were fully above the statewide average were noted as high performers.

Additionally, the Top 5 performers are identified at the medical group level and Top 15 performers are identified at the clinic level.

TOP = Top 15 clinics or Top 5 medical groups, as long as the clinic or medical group meets the Above rating criteria. For measures that have less than 10 reportable medical groups, there is no TOP assignment.

ABOVE = Medical group/clinic rate and 95 percent confidence intervals are fully above the statewide average rate.

AVERAGE = Medical group/clinics with 95 percent confidence intervals that crosses the statewide average rate.

BELOW = Medical group/clinic rate and 95 percent confidence intervals are fully below the statewide average rate.

Identifying Medical Groups and Clinics with Biggest Improvements

For each measure, individual medical group and clinic rates during report year 2015 were compared with their rates during report year 2014, calculating an absolute percentage point difference. Medical groups and clinics with the largest percentage point increases were identified.

Medical Group and Clinic Performance Over Time (Three Years)

This analysis was done to determine patterns of medical group and clinic performance over time per measure. Patterns were reviewed for the three reporting years (2013, 2014 and 2015).

The percent and number of medical groups were reported for each of the following patterns of rate changes over the past three years for each measure:

• Consistently improved: Medical groups with more than a two percentage point increase between each consecutive year.

• Relatively stable: Medical groups that had no more than a two percentage point increase or decrease between each consecutive year (-2 percent – +2 percent).

• Consistently decreased: Medical groups with more than a two percentage point decrease between each consecutive year.

• Variable performance (with an improvement or with a decline): Medical groups with an up/down pattern that was not consistent and did not fall into one of the other categories.

7.2.4.1. Confidence intervals

Confidence intervals using the method of Agresti and Coull

The Wilson method for calculating confidence intervals for proportions (introduced by Wilson (1927), recommended by Brown, Cai and DasGupta (2001) and Agresti and Coull (1998)) is based on inverting the hypothesis test given in Section 7.2.4. That is, solve for the two values of p_0 (say, p_{upper} and p_{lower}) that result from setting $z = z_{1-\alpha/2}$ and solving for $p_0 = p_{upper}$, and then setting $z = z_{\alpha/2}$ and solving for $p_0 = p_{lower}$. (Here, as in Section 7.2.4, $z_{\alpha/2}$ denotes the variate value from the standard normal distribution such that the area to the left of the value is $\alpha/2$.) Although solving for the two values of p_0 might sound complicated, the appropriate expressions can be obtained by straightforward but slightly tedious algebra. Such algebraic manipulation isn't necessary, however, as the appropriate expressions are given in various sources. Specifically, we have

Formulas for the confidence intervals

$$\text{U.L.} = \frac{\hat{p} + \frac{z^2_{1-\alpha/2}}{2n} + z_{1-\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z^2_{1-\alpha/2}}{4n^2}}}{1 + \frac{z^2_{1-\alpha/2}}{n}}$$

$$\text{L.L.} = \frac{\hat{p} + \frac{z^2_{\alpha/2}}{2n} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z^2_{\alpha/2}}{4n^2}}}{1 + \frac{z^2_{\alpha/2}}{n}}.$$

The Wilson method for calculating confidence intervals for all clinic rates and statewide rates.

www.itl.nist.gov/div898/handbook/prc/section2/prc241.htm

2b5.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

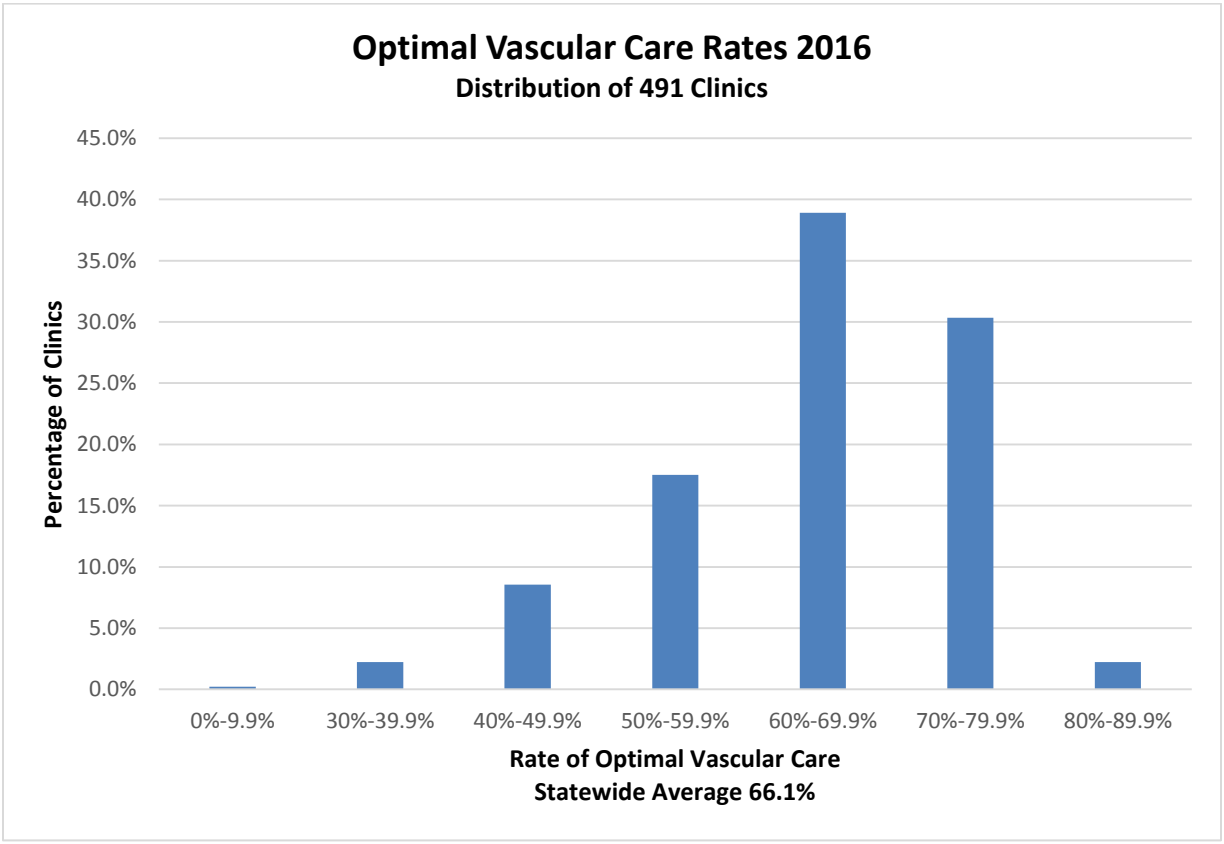


TABLE 12: STATEWIDE RATE FOR OPTIMAL VASCULAR CARE

	Statewide Average (Weighted)	95% CI	Numerator (Patients who met treatment goals)	Denominator (Patients Sampled)	Total Eligible
Optimal Vascular Care	69.3%	69.0%-69.6%	71,196	102,654	103,006

APPENDIX

Highest Performers in 2015 by Medical Group Type for Clinical Measures

The following **tables 33-37** list the high-performing medical groups by type based on the clinical measures on which they were reported. These groups had above-average rates on most clinical measures.

Seven primary care medical groups achieved rates that were above average more frequently than others. Each of these medical groups achieved above-average rates on at least half of the clinical measures on which they were reported.

TABLE 33: HIGH PERFORMING MEDICAL GROUPS IN 2015 - **PRIMARY CARE**

Measure	Health Partners Clinics 16 out of 20	Park Nicollet Health Services 15 out of 20	Allina Health 13 out of 20	Stillwater Medical Group 13 out of 20	Mankato Clinic, Ltd. 11 out of 20	Essentia Health - East Region 10 out of 20	Mayo Clinic 10 out of 20
ADHD	●						
Adolescent Immunizations	●	●	●	●	●	●	
Breast Cancer Screening	●	●	●		●	●	●
Bronchitis	●	●		●			
Childhood Immunization Status (Combo 3)	●				●		●
Chlamydia Screening	●	●	●	●			
Colorectal Cancer Screening	●	●	●	●	●		●
Controlling High Blood Pressure		●	●	●		●	
COPD	●	●					●
Depression Remission at Six Months		●	●			●	●
Depression Remission at Twelve Months			●		●	●	●
Maternity Care: Primary C-Section Rate			●				●
Pharyngitis	●	●	●	●	●		
Optimal Asthma Control-Children	●	●		●	●	●	
Optimal Asthma Control-Adults	●	●	●	●		●	
Optimal Diabetes Care	●	●		●	●	●	
Optimal Vascular Care	●	●	●	●	●		
URI	●	●	●	●		●	●
Pediatric Mental Health Screening	●		●	●	●		●
Pediatric Overweight Counseling	●	●		●	●	●	●

● = Medical group rate and CI fully above average
Blank = Measure reported but rate was average or below average.

Disparities Report:
Significant gaps still exist for patients under MN Health Care Programs versus all other payers

Year	MHCP Rate	MHCP CI (U/L)*	MHCP Denominator (Patients Sampled)	Other Purchaser Rate	Difference (Other-MHCP)
2015^	58.3%	57.3% - 59.3%	10,152	72.2%	13.9%**

* A confidence interval gives an estimated range of values calculated from a given sample of data. A 95 percent confidence interval implies a 95 percent level of confidence that the interval includes the true mean or parameter.
** Denotes a statistically significant difference.
^ LDL component removed from measure calculation.

2b5.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)
Measure identifies both opportunity for improvement in outcomes and processes to reduce risk of long term complications for patients with diabetes and identifies meaningful differences among providers.

2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS
Note: *Applies to all component measures, unless already endorsed or are being submitted for individual endorsement.* ***If only one set of specifications for each component, this section can be skipped.***

Note*: This item is directed to measures that are risk-adjusted (with or without SDS factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without SDS factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b6.1. Describe the method of testing conducted to demonstrate comparability of performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

2b6.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

2b6.3. What is your interpretation of the results in terms of demonstrating comparability of performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted?*)

2b7. MISSING DATA ANALYSIS AND MINIMIZING BIAS
Note: *Applies to the overall composite measure.*
2b7.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)
For this patient level all-or-none composite measure, elements missing from any component (e.g. visit but no blood pressure during the measurement year) are counted as a numerator component fail and therefore the patient would be accounted for and remain in the denominator.

2b7.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

The impact of missing data on measure calculations is minimal. For 2013 dates of service on over 104,395 ischemic vascular patients submitted for rate calculation two variables were considered 1) with in the appropriate measurement timeframe and 2) valid values submitted:

Variable	Within measure period	Invalid values
Blood Pressure	99.8%	0.03%
Tobacco Status documented	99.4%	0.0%
Aspirin or anti-platelets *	96.7%	
Statin **	94.3%	

* had documented aspirin or anti-platelet in the measurement year or the date of a valid contraindication
** had documented statin in the measurement year or the date of a valid contraindication

2b7.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

The impact of missing data on measure calculations is minimal. Patients with missing data are not excluded from the measure. Elements missing from any component are counted as a numerator component fail and remain in the denominator.

2d. EMPIRICAL ANALYSIS TO SUPPORT COMPOSITE CONSTRUCTION APPROACH

Note: *If empirical analyses do not provide adequate results—or are not conducted—justification must be provided and accepted in order to meet the must-pass criterion of Scientific Acceptability of Measure Properties. Each of the following questions has instructions if there is no empirical analysis.*

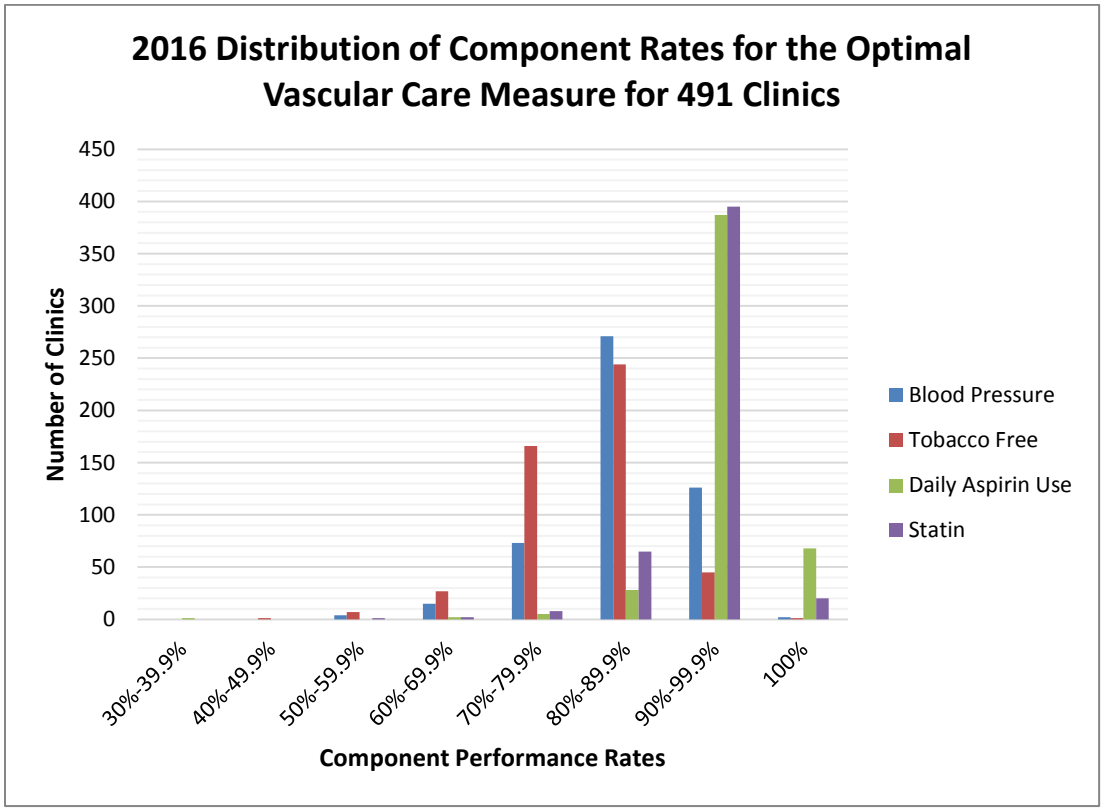
2d1. Empirical analysis demonstrating that the component measures fit the quality construct, add value to the overall composite, and achieve the object of parsimony to the extent possible.

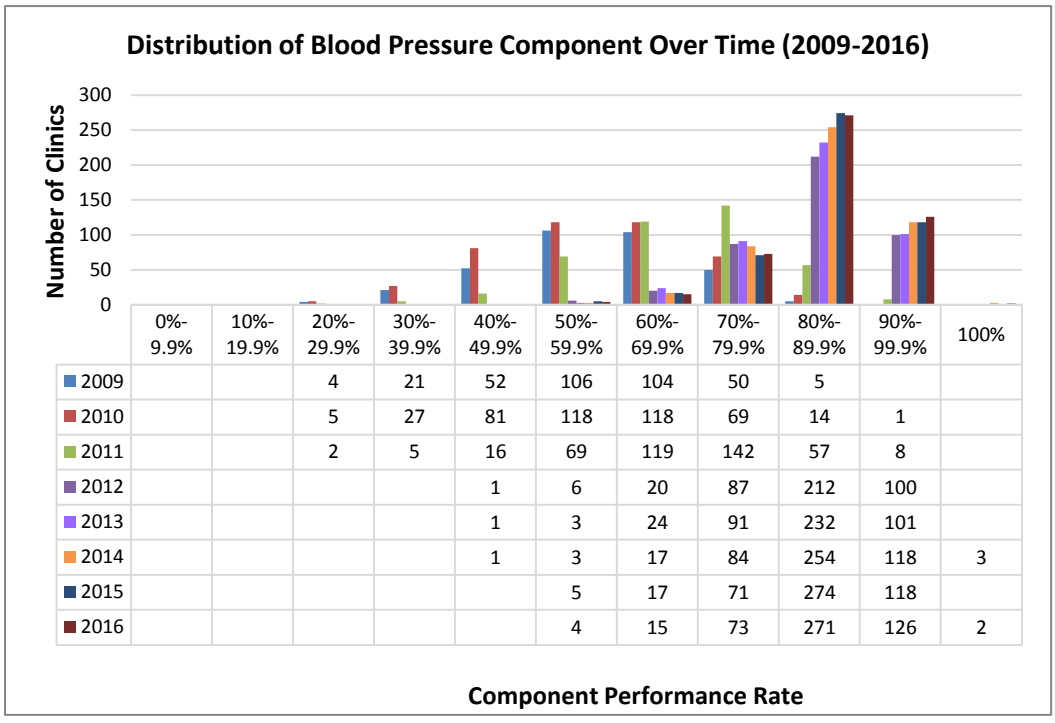
This composite measure is a patient level all-or-none composite in which the desired goal is for the patient is to achieve multiple intermediate physiological clinical outcome and medication use targets to best reduce their overall risk of developing long term complications or additional cardiovascular events. Reducing modifiable risks was the reason why this measure was developed. The components of this measure include blood pressure control, being tobacco-free, appropriate use of statins and daily aspirin or anti-platelet.

Achieving the intermediate physiological outcome targets related to blood pressure and glycemic control in addition being tobacco free and use of daily aspirin and statins where appropriate are the ischemic vascular disease patient's best mechanisms of avoiding or postponing long term complications associated with this chronic condition which affects millions of Americans. Measuring providers separately on individual targets is not as patient centric as a measure that seeks to reduce multiple risk factors for each patient. Patients are more likely to reduce their overall risk and maximize health outcomes by achieving several intermediate physiological targets.

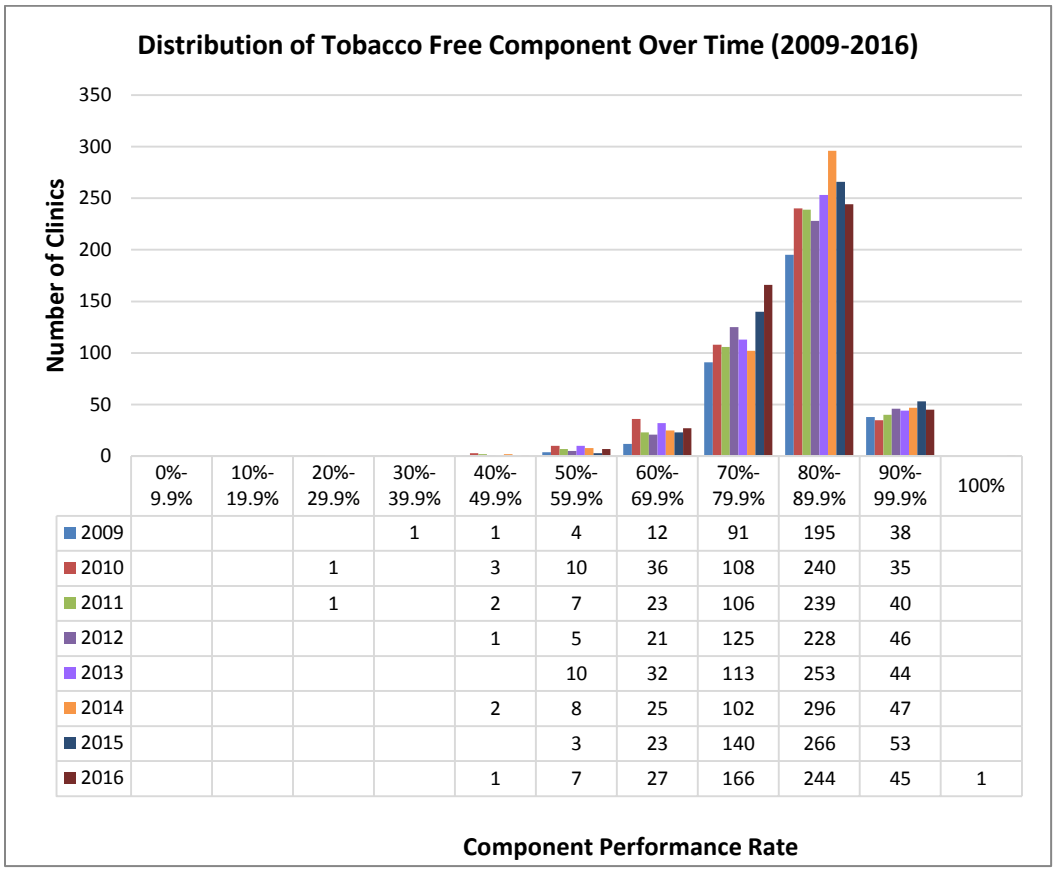
The components of this patient level all-or-none composite measure, though they can be analyzed as individual components especially for purposes of understanding opportunities within the composite measure, are treated as a whole. There is no weighting of the components; it is an all-or-none measure.

This graph demonstrates how clinics are performing for each component of the measure:

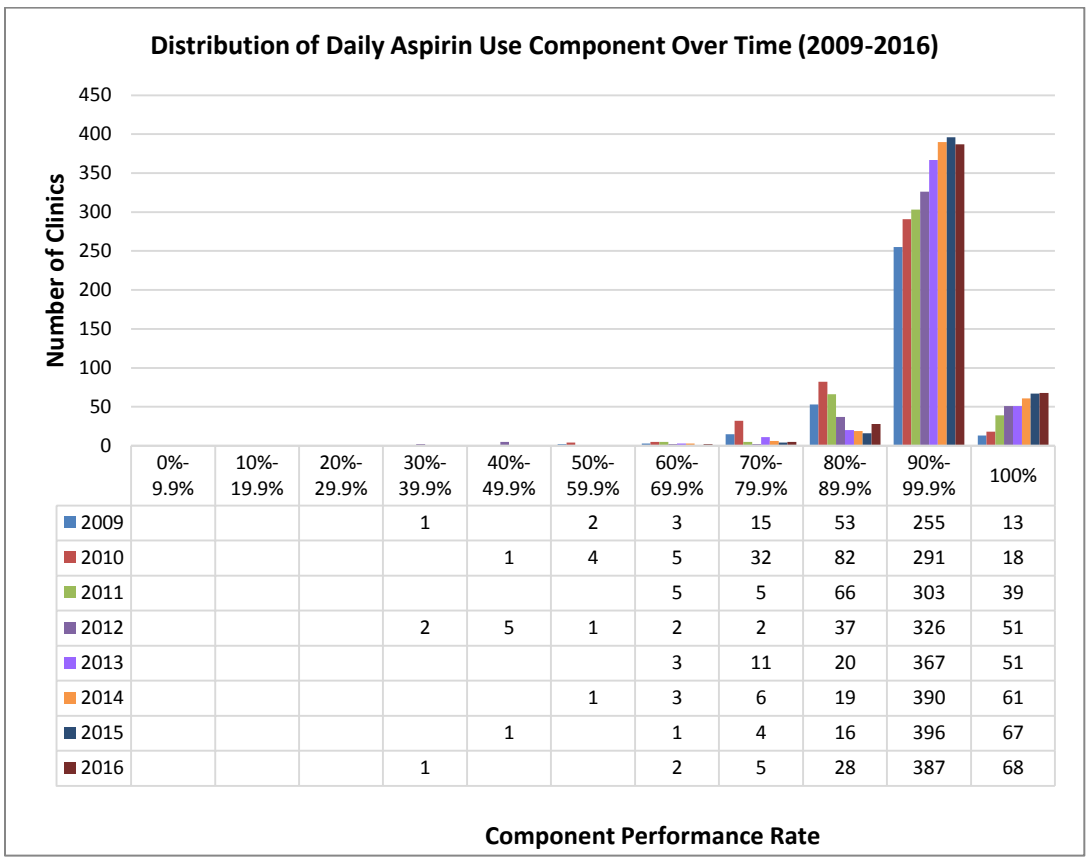




	0%-9.9%	10%-19.9%	20%-29.9%	30%-39.9%	40%-49.9%	50%-59.9%	60%-69.9%	70%-79.9%	80%-89.9%	90%-99.9%	100%
2009			4	21	52	106	104	50	5		
2010			5	27	81	118	118	69	14	1	
2011			2	5	16	69	119	142	57	8	
2012					1	6	20	87	212	100	
2013					1	3	24	91	232	101	
2014					1	3	17	84	254	118	3
2015						5	17	71	274	118	
2016						4	15	73	271	126	2



	0%-9.9%	10%-19.9%	20%-29.9%	30%-39.9%	40%-49.9%	50%-59.9%	60%-69.9%	70%-79.9%	80%-89.9%	90%-99.9%	100%
2009				1	1	4	12	91	195	38	
2010			1		3	10	36	108	240	35	
2011			1		2	7	23	106	239	40	
2012					1	5	21	125	228	46	
2013						10	32	113	253	44	
2014					2	8	25	102	296	47	
2015						3	23	140	266	53	
2016					1	7	27	166	244	45	1



	0%-9.9%	10%-19.9%	20%-29.9%	30%-39.9%	40%-49.9%	50%-59.9%	60%-69.9%	70%-79.9%	80%-89.9%	90%-99.9%	100%
2009				1		2	3	15	53	255	13
2010					1	4	5	32	82	291	18
2011							5	5	66	303	39
2012				2	5	1	2	2	37	326	51
2013							3	11	20	367	51
2014						1	3	6	19	390	61
2015					1		1	4	16	396	67
2016				1			2	5	28	387	68

Proportion of how many patients are meeting component targets (2016)

Category	# of Patients	Proportion
Daily Aspirin Use alone	100,918	96.7%
Statin alone	98,653	94.5%
Blood Pressure alone	88,770	85.0%
Tobacco Free alone	86,680	83.0%
Daily Aspirin Use + Statin	96,137	92.1%
Blood Pressure + Daily Aspirin Use	86,037	82.4%
Blood Pressure + Statin	84,147	80.6%
Tobacco + Daily Aspirin Use	84,000	80.5%
Tobacco + Statin	82,091	78.6%
Blood Pressure + Tobacco	74,125	71.0%
Blood Pressure + Daily Aspirin Use + Statin	82,144	78.7%
Tobacco Free + Daily Aspirin Use + Statin	80,159	76.8%
Blood Pressure + Tobacco Free + Daily Aspirin Use	71,983	69.0%
Blood Pressure + Tobacco Free + Statin	70,418	67.4%
All 4 components (statewide average)	69,026	66.1%

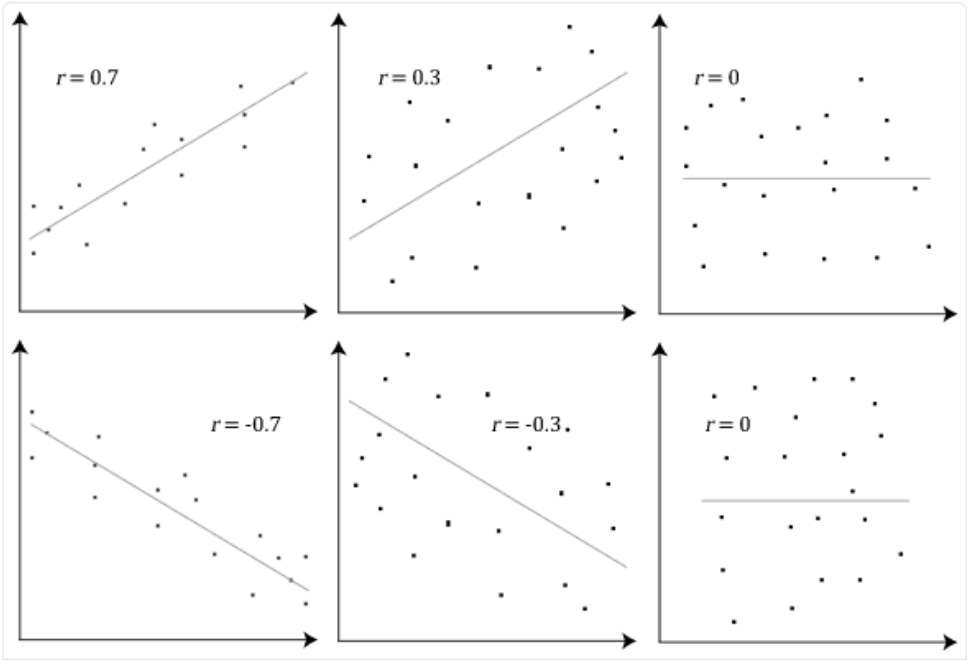
2d1.1 Describe the method used (*describe the steps—do not just name a method; what statistical analysis was used; if no empirical analysis, provide justification)*

https://statistics.laerd.com/statistical-guides/pearson-correlation-coefficient-statistical-guide.php

The Pearson product-moment correlation coefficient (or Pearson correlation coefficient, for short) is a measure of the strength of a linear association between two variables and is denoted by *r*. Basically, a Pearson product-moment correlation attempts to draw a line of best fit through the data of two variables, and the Pearson correlation coefficient, *r*, indicates how far away all these data points are to this line of best fit (how well the data points fit this new model/line of best fit).

The Pearson correlation coefficient, r, can take a range of values from +1 to -1. A value of 0 indicates that there is no association between the two variables. A value greater than 0 indicates a positive association; that is, as the value of one variable increases, so does the value of the other variable. A value less than 0 indicates a negative association; that is, as the value of one variable increases, the value of the other variable decreases.

The stronger the association of the two variables, the closer the Pearson correlation coefficient, r, will be to either +1 or -1 depending on whether the relationship is positive or negative, respectively. Achieving a value of +1 or -1 means that all your data points are included on the line of best fit - there are no data points that show any variation away from this line. Values for r between +1 and -1 (for example, r = 0.8 or -0.4) indicate that there is variation around the line of best fit. The closer the value of r to 0 the greater the variation around the line of best fit. Different relationships and their correlation coefficients are shown in the diagram below:



2d1.2. What were the statistical results obtained from the analysis of the components? (e.g., *correlations, contribution of each component to the composite score, etc.*; *<u>if no empirical analysis</u>, identify the components that were considered and the pros and cons of each*)

Optimal Vascular Care and Blood Pressure

Correlation Analysis

The CORR Procedure

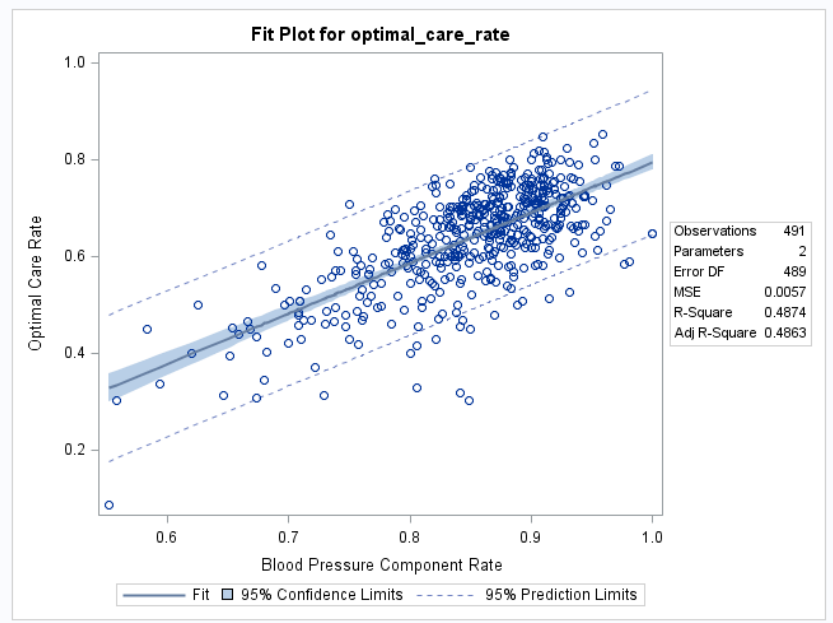
1 With Variables:	blood_pressure_component_rate
1 Variables:	optimal_care_rate

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
blood_pressure_component_rate	491	0.85048	0.07079	417.58775	0.55172	1.00000	Blood Pressure Component Rate
optimal_care_rate	491	0.63919	0.10564	313.84231	0.08621	0.85338	Optimal Care Rate

Pearson Correlation Coefficients, N = 491
Prob > |r| under H0: Rho=0

	optimal_care_rate
blood_pressure_component_rate	0.69813
Blood Pressure Component Rate	<.0001



Optimal Vascular Care and Tobacco Free

Correlation Analysis

The CORR Procedure

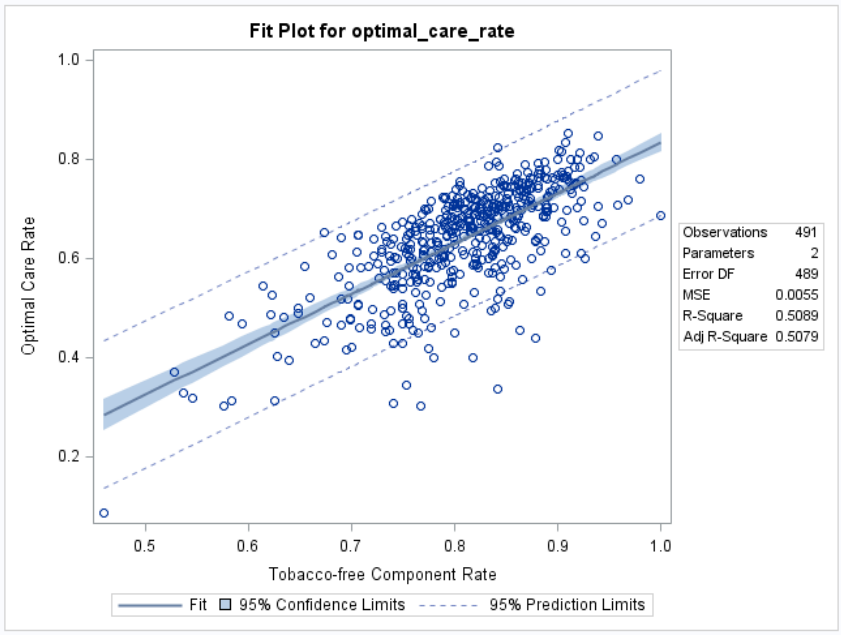
1 With Variables:	tobacco_free_component_rate
1 Variables:	optimal_care_rate

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
tobacco_free_component_rate	491	0.80901	0.07427	397.22177	0.45977	1.00000	Tobacco-free Component Rate
optimal_care_rate	491	0.63919	0.10564	313.84231	0.08621	0.85338	Optimal Care Rate

Pearson Correlation Coefficients, N = 491
Prob > |r| under H0: Rho=0

	optimal_care_rate
tobacco_free_component_rate	0.71336
Tobacco-free Component Rate	<.0001



Optimal Vascular Care and Daily Aspirin Use

Correlation Analysis

The CORR Procedure

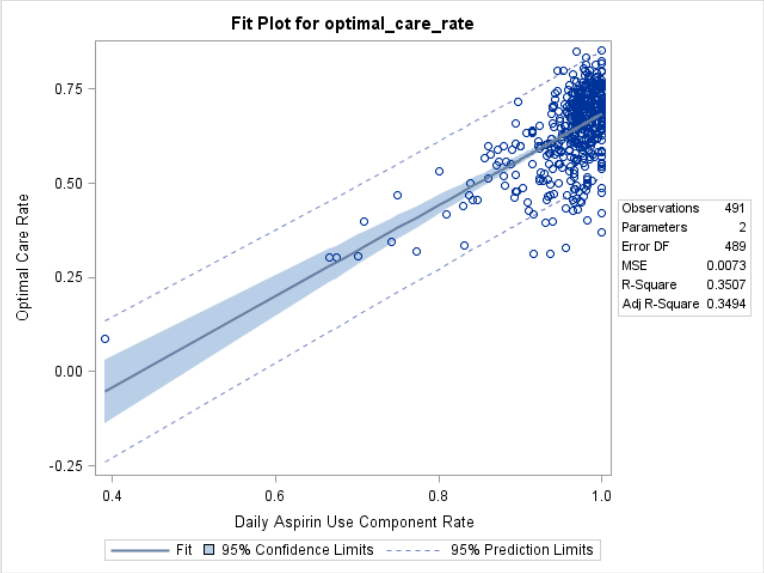
1 With Variables:	daily_aspirin_use_component_rate
1 Variables:	optimal_care_rate

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
daily_aspirin_use_component_rate	491	0.96271	0.05167	472.68933	0.39080	1.00000	Daily Aspirin Use Component Rate
optimal_care_rate	491	0.63919	0.10564	313.84231	0.08621	0.85338	Optimal Care Rate

Pearson Correlation Coefficients, N = 491
Prob > |r| under H0: Rho=0

	optimal_care_rate
daily_aspirin_use_component_rate	0.59223
Daily Aspirin Use Component Rate	<.0001



Optimal Vascular Care and Statin Use

Correlation Analysis

The CORR Procedure

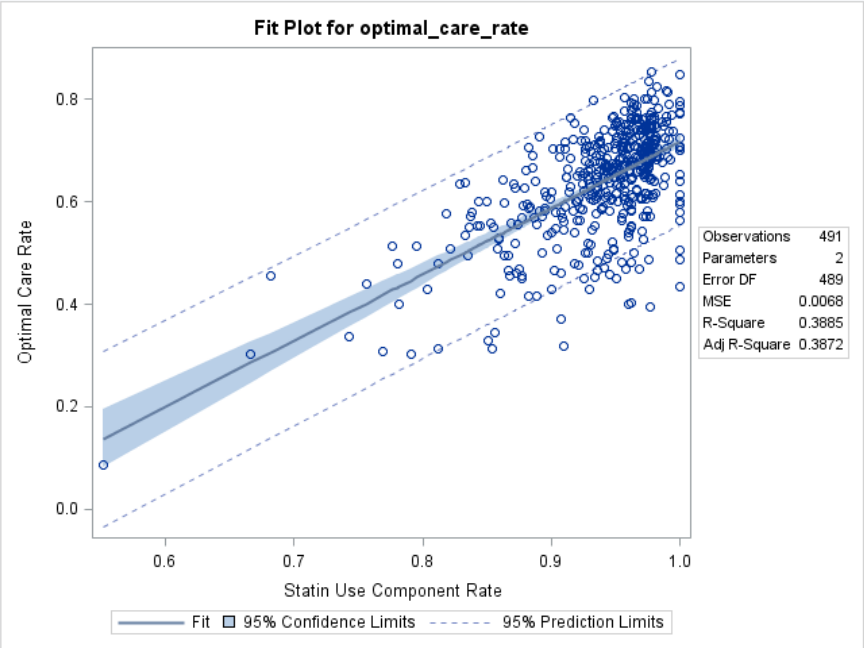
1 With Variables:	statin_use_component_rate
1 Variables:	optimal_care_rate

Simple Statistics

Variable	N	Mean	Std Dev	Sum	Minimum	Maximum	Label
statin_use_component_rate	491	0.93973	0.05086	461.40755	0.55172	1.00000	Statin Use Component Rate
optimal_care_rate	491	0.63919	0.10564	313.84231	0.08621	0.85338	Optimal Care Rate

Pearson Correlation Coefficients, N = 491
Prob > |r| under H0: Rho=0

	optimal_care_rate
statin_use_component_rate	0.62327
Statin Use Component Rate	<.0001



2d1.3. What is your interpretation of the results in terms of demonstrating that the components included in the composite are consistent with the described quality construct and add value to the overall composite? (i.*e., what do the results mean in terms of supporting inclusion of the components; if no empirical analysis, provide rationale for the components that were selected)*

Practices in Minnesota demonstrate high compliance with and implementation of clinical guidelines for prescribing/ordering medications for patients with ischemic vascular disease that reduce their risk for future events or long term complications. 2009 to 2011 demonstrated more variability in the daily aspirin/ anti-platelet component, but as groups integrated this into their practices the component rates increased and held steady. Aspirin/ anti-platelet rates are not necessarily as high across the country. Paerkh A.K. et al cite rates as low as 34.8% in primary care settings. [Aspirin in the Secondary Prevention of Cardiovascular Disease, NEJM Jan 2013]. While all components demonstrate opportunity for improvement at a clinic level, blood pressure control and tobacco free components demonstrate more variability, opportunity for improvement and impact ability to achieve all four components.

Pearson Correlation Analysis

Components as compared to the composite Optimal Care Rate demonstrate a strong correlation with the following Pearson r coefficient values: Blood Pressure at 0.69813, Tobacco-free at 0.71336, Aspirin or Anti-platelet Use at 0.59223 and Statin Use at 0.62327.

2d2. Empirical analysis demonstrating that the aggregations and weighting rules are consistent with the quality construct and achieve the objective of simplicity to the extent possible

2d2.1 Describe the method used (*describe the steps—do not just name a method; what statistical analysis was used; if no empirical analysis, provide justification)*

2d2.2. What were the statistical results obtained from the analysis of the aggregation and weighting rules? (e.g., *results of sensitivity analysis of effect of different aggregations and/or weighting rules; if no empirical analysis, identify the aggregation and weighting rules that were considered and the pros and cons of each*)

2d2.3. What is your interpretation of the results in terms of demonstrating the aggregation and weighting rules are consistent with the described quality construct? (i.*e., what do the results mean in terms of supporting the selected rules for aggregation and weighting; if no empirical analysis, provide rationale for the selected rules for aggregation and*

weighting)

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. **Data Elements Generated as Byproduct of Care Processes.**

generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition,
Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)
If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. **To what extent are the specified data elements available electronically in defined fields?** *(i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields)*

ALL data elements are in defined fields in electronic health records (EHRs)

3b.2. **If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources.**

3b.3. **If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL.**

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.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 and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.**

IF a PRO-PM, consider implications for both individuals providing PROM data (patients, service recipients, respondents) and those whose performance is being measured.

Over the last several years we have learned the following:

1. Data Submission- Providing data collection software for medical groups wishing to submit data was not always the best and most efficient way of collecting data. As electronic health records use becomes more pervasive in our state, providing templates of data file submissions proved to be more efficient.
2. Specifications- Detailed specifications with instructions on how to handle most situations (e.g. detailed instructions on blood pressure values) has been valuable to medical groups, increased data accuracy and resulted in successful and accurate data submission by all the practices in MN.
3. Audit- Audit methods have ensured the accuracy of our data and we are able to successfully compare providers because everyone is pulling their data the same way and subject to the same rules.
4. Confidentiality- Patient confidentiality has been addressed by numerous mechanisms. MNCM only receives the patient level information needed to calculate the rates, determine eligibility for inclusion in the measure and support the administration of pay

for performance programs. The PHI submitted is minimal and the data is protected by 1) password protection with password only available to the medical group submitting data, 2) file upload process is encrypted as data is transferred and 3) Data is stored on a separate secure server and meets all HIPAA protection rules.
5. Electronic Medical Record- It is easier for groups that have an electronic medical record to submit data and to submit their full population of patients, however many groups with paper chart systems can successfully submit their sample.
6. Acceptance of Data- Vast improvement in terms of sample sizes and timeliness of the data submitted by medical groups six weeks after the end of the measurement year as compared to prior method of health plan's samples and the results over a year old. Providers are more accepting of the results as compared to previous methods of pooling health plan samples.
7. Data Collection Burden- We have learned that for additional future measures we will need to stagger the data collection time frames and submission deadlines as to not burden the medical groups in terms of abstraction/ extraction (e.g. can't always have a measurement period Jan 1st to Dec 31st reported the second week of February, may need to consider July 1st to June 30th with data submission in August)
8. Health Plans: pay for performance and the inclusion of measures within contracts significantly impacts the number of groups participating in each measure (Diabetes, Ischemic Vascular, and Depression)

3c.2. **Describe any fees, licensing, or other requirements to use any aspect of the measure as specified** *(e.g., value/code set, risk model, programming code, algorithm)*.
There are no fees associated with participation and submitting data for this measure. Results are available to 1) all data submitters within the HIPAA secure MNCM data portal and 2) to the public on our consumer facing website MN Health Scores at www.mnhealthscores.org and 3) annual health care quality report on our corporate website at www.mncm.org. There are costs to the medical groups in terms of extract programs or abstraction to submit patient level clinical information for rate calculation.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. **Current and Planned Use**

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Planned	Current Use (for current use provide URL)
Quality Improvement (Internal to the specific organization)	Public Reporting MN HealthScores http://www.mnhealthscores.org/clinic-measure-detail/vascular-disease/#/results Payment Program Minnesota Bridges to Excellence and MN Quality Incentive Payment System http://mnhealthactiongroup.org/taking-action/innovation/care-system-quality-and-performance/bridges-to-excellence/ Quality Incentive Payment System (QIPS) http://www.health.state.mn.us/healthreform/measurement/qips.html Regulatory and Accreditation Programs Minnesota Statewide Quality Reporting and Measurement System (SQRMS) http://www.health.state.mn.us/healthreform/measurement/index.html

	Quality Improvement with Benchmarking (external benchmarking to multiple organizations) Minnesota Department of Human Services (DHS) http://mn.gov/health-reform/topics/prevention/health-disparities/index.jsp

4a.1. For each CURRENT use, checked above, provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included

MN Community Measurement- MN HealthScores Website
Public Reporting consumer-facing website
All primary care, multi-specialty clinics with cardiology services and cardiology clinics in Minnesota (mandatory) and bordering communities (voluntary)
111 Medical groups representing 671 clinic sites; 2015 dates of service 104,494 patients with ischemic vascular disease.

MNCM Health Care Quality, Equity and Disparity Annual Reports
Public Reporting: Hard-copy reports (pdf) highlighting top performers, most improved
111 Medical groups representing 671 clinic sites; 2015 dates of service 104,494 patients with ischemic vascular disease.

MN Bridges to Excellence and MN Quality Incentive Payment System
Pay for Performance Program for top performers and those attaining improvement goals/ benchmarks. All clinics participating in MNCM's data portal who submit full population are eligible for inclusion in this program. Annual recognition event and rewards distributed.

Quality Incentive Payment System (QIPS)
The Minnesota Quality Incentive Payment System (QIPS) is a statewide pay-for-performance system for physician clinics and hospitals. It is built on the measures of the Statewide Quality Reporting and Measurement System (Quality Reporting System), Minnesota's standardized set of quality measures for health care providers. The Minnesota Department of Health (MDH) updates QIPS on a yearly basis. The three clinic measures included in the system are Optimal Diabetes Care, Optimal Vascular Care and Depression Remission at Six Months. The ten hospital measures relate to patient satisfaction. The system rewards providers for two types of accomplishment: (1) achieving absolute performance benchmarks or (2) improvements in performance over time. In 2014, MMB and DHS paid nearly $500,000 in incentive payments to providers in 259 clinics that achieved benchmarks or significantly improved care for diabetes, vascular disease, and depression.

MN Department of Health- Statewide Quality Reporting and Measurement System
Based on 2008 health reform state legislation; this program requires mandatory submission of data from Minnesota physician clinics that have provider specialties that are applicable to the measured population. For the Optimal Vascular Control Measure: family medicine, general practice, internal medicine, geriatric medicine and cardiology.

Minnesota Department of Human Services (DHS)
This competitive grant program is administered by the Minnesota Department of Health's Office of Minority and Multicultural Health and provides funds to close the gap in the health status of African Americans/Africans, American Indians, Asian Americans, and Hispanics/Latinos in Minnesota as compared to Whites in the following priority health areas: Breast and cervical cancer screening, diabetes, heart disease and stroke, HIV/AIDS and sexually-transmitted diseases, immunizations, infant mortality, teen pregnancy prevention and unintentional injury and violence.

4a.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

4a.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)

4b. Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b.1. Progress on Improvement. (Not required for initial endorsement unless available.)

Performance results on this measure (current and over time) should be provided in 1b.2 and 1b.4. Discuss:

- **Progress (trends in performance results, number and percentage of people receiving high-quality healthcare)**
- **Geographic area and number and percentage of accountable entities and patients included**

Although the state-wide average, reflecting over 104,000 patients, may appear to be making only small incremental improvements year to year, please note the increasing number of numerator cases. Between report years 2013 and 2014; prior to measure redesign, 6,719 more patients achieved optimal control targets than the previous year.

For 2016 (2015 dates of service), 66.1% of the patients met all four component targets in the composite measure and were considered optimally managed. 111 medical groups representing 671 physician clinics and 104,494 patients with IVD in Minnesota and neighboring communities submitted data for this measure. Of the clinics that were reportable (patient n >= 30), there was a wide range of variability with the lowest scoring clinic at 8.6% and the highest scoring clinic at 85.3%.

The trends for this measure are as follows:

Report Year	Rate	Patients (Den)	Numerator	Eligible	% submit/eligible
2007	38.9%	4,662	1,595	11,740	39.7%
2008	32.6%	36,126	11,997	54,708	66.0%
2009	33.8%	46,779	16,529	80,907	57.8%
2010	33.8%	63,241	21,589	95,791	66.0%
2011*	39.7%	66,910	27,083	96,270	69.5%
2012	49.4%	78,886	39,242	95,482	82.6%
2013	48.5%	87,345	42,689	93,761	93.1%
2014	50.0%	98,803	49,408	99,550	99.2%
2015**	69.3%	102,654	71,196	103,006	99.7%
2016***	66.1%	104,395	69,026	104,494	99.9%

* Blood pressure component target change based on evidence/ guidelines from < 130/80 to < 140/90
** Cholesterol management component suppressed during re-design
*** Cholesterol management component change from LDL < 100 to appropriate statin use

Individual rates of the components are as follows:
Blood Pressure <140/90 = 85.0%
Statin Use = 95.2%
Daily Aspirin Use = 96.7%
Tobacco Non-user = 83.0%

Please note that while the all-or-none composite measure is considered to be the gold standard, reflecting best patient outcomes, the individual components may be measured as well. This is particularly helpful in quality improvement efforts to better understand where opportunities exist in moving the patients toward achieving all of the desired outcomes. Please refer to the additional numerator logic provided for each component.

Trend over time by Component and Report Year

	2009	2010	2011	2012	2013	2014	2015	2016	
BP <140/90	-	-	-	-	84.0%	84.1%	84.9%	85.2%	85.0%
Aspirin Use	92.5%	91.9%	94.2%	94.7%	96.5%	96.6%	96.6%	96.7%	
Tobacco Free	82.4%	81.2%	82.7%	82.6%	82.9%	84.1%	83.5%	83.0%	

4b.2. If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4c. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4c.1. Were any unintended negative consequences to individuals or populations identified during testing; OR has evidence of unintended negative consequences to individuals or populations been reported since implementation? If so, identify the negative unintended consequences and describe how benefits outweigh them or actions taken to mitigate them.

No unintended consequences identified during the testing, implementation and ongoing review of this measure.

MN Community Measurement has modeled the direct data submission to minimize inaccuracies, errors and unintended consequences. All groups participating sign a terms of use agreement that delineates the group's responsibilities for submission of data and consequences for not participating in good faith. Additionally all groups sign a Business Associate Agreement that outlines the use of the data. Denominator certification prior to any data collection ensures that groups are following the specifications and correctly identifying their population and serves as a point of correction prior to the expenditure of resources for data collection. Groups provide documentation of cases that are excluded and this is reviewed by MNCM staff prior to approval of the data submission. Extensive audit processes also support the data's accuracy. After data submission, in person validation audits are conducted comparing the submission to the patient´s medical record using NCQA´s 8 and 30 rule for audit requiring a 90% accuracy rate. Groups are only allowed three patient records with error out of 30 reviewed in order to achieve 90%. Audits are conducted in the following instances: 1) a random sample of clinics with prior successful submission, 2) for all groups who are new to the submission process, 3) a group who has had a change in system or process (e.g. went from paper charts to EMR) since the last submission or 4) any group with a history of prior unsuccessful audit. It has been our experience that the post submission audits have identified both issues with data extraction programming from an EMR and abstraction errors when data is collected from the chart. Groups have been amenable to remedy plans, resubmission and re-audit.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

Yes

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

0067 : Chronic Stable Coronary Artery Disease: Antiplatelet Therapy
0068 : Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antiplatelet
0073 : Ischemic Vascular Disease (IVD): Blood Pressure Control
0543 : Adherence to Statin Therapy for Individuals with Cardiovascular Disease

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

There is a new measure submitted for endorsement during this phase; #2763 Ischemic Vascular Disease Care: All or None Outcome Measure-Optimal Control by the Wisconsin Collaborative for Healthcare Quality (WCHQ) that is directly competing with this measure.

5a. Harmonization

The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):
Are the measure specifications completely harmonized?

No

5a.2. **If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.**
There are some differences noted in the denominator definitions, source data and settings of care. #0068 Ischemic Vascular Disease (IVD): Use of Aspirin or Another Antiplatelet AND #0073 Ischemic Vascular Disease (IVD): Blood Pressure Control are most closely related to the components of our measure, however this measure focuses on the inpatient setting and only patients discharged with acute myocardial infarction, coronary bypass graft or percutaneous coronary interventions. #0067 Chronic Stable Coronary Artery Disease: Antiplatelet Therapy focuses only on patients with coronary artery disease; however from specifications available through QPS not able to compare diagnosis code definitions. This measure, #0076 Optimal Vascular Care is more inclusive with a denominator definition of ischemic vascular disease (atherosclerosis of coronary and peripheral arteries) #0543 Adherence to statin therapy for individuals with cardiovascular disease. This medication claims based measure's denominator is more aligned with our intent (coronary, cerebrovascular and peripheral artery disease), however endorsement was removed in 2015.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. **If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):**
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)
There are other similar measures that address three of the four components separately, but no currently endorsed measure exists that is a patient level all-or-none composite measure.

0076 Optimal Vascular Care is superior to the newly submitted measure for consideration because its measure construct additionally includes:
* contraindications and exceptions to statin use
* risk adjustment; actual and expected rates reported
* allowable exclusions for potentially frail older adults age 65 to 75 (hospice or palliative services, nursing home, death)

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment:

Contact Information

Co.1 **Measure Steward (Intellectual Property Owner):** MN Community Measurement
Co.2 **Point of Contact:** Jasmine, Larson, jlarson@mncm.org, 612-746-4514-
Co.3 **Measure Developer if different from Measure Steward:** MN Community Measurement
Co.4 **Point of Contact:** Jasmine, Larson, jlarson@mncm.org, 612-746-4514-

Additional Information

Ad.1 **Workgroup/Expert Panel involved in measure development**
Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
#0076 Optimal Vascular Care was originally stewarded by HealthPartners and received its initial NQF endorsement in 2009. HealthPartners turned over the stewardship of the four component patient level all-or-none composite measure to MN Community Measurement (MNCM) in 2010.

In response to guideline changes over the year's MNCM's Measurement and Reporting Committee has authorized ad-hoc measure development groups to re-design or modify measure targets based on significant changes in the evidence and guidelines.
Blood pressure numerator targets were modified in 2010 using an expedited process following the Institute for Clinical Systems guideline publication to make changes to both the Optimal Diabetes Care and Optimal Vascular Care measures.

Original Workgroup included:
Beth Averbeck, MD HeathPartners
Rich Bergenstal, MD International Diabetes Center Park Nicollet
Barry Bershow, MD, Fairview Health Services
John Fredrick, MD Preferred One
Diane Mayberry, MN Community Measurement
Victor Montori, MD Mayo Clinic
Mark Nyman, MD Mayo Clinic
Gene Ollila, MD Allina Medical Clinic
Collette Pitzen, MN Community Measurement
Kari Retzer, ICSI Facilitator for Diabetes Guideline
JoAnn Sperl-Hillen, MD HealthPartners
Linda Walling, MD, HealthEast

In September 2013, MARC requested a diabetes measure development work group ad-hoc review of the cholesterol component based on ongoing comments received to consider modification of the LDL component to "LDL < 100 or patient is on a statin". As work group member recruitment proceeded, the advent of the long awaited updated guidelines necessitated a more extensive consideration for revision of the cholesterol/ lipid target component for both the Optimal Diabetes Care and Optimal Vascular Care measures

The measure development work group met to discuss the new guidelines and determine the future direction for the cholesterol/ lipid component of MNCM's diabetes measure. After thoughtful consideration of new guidelines that focus on statin use and discourage targeting treatment to achieve certain cholesterol levels, the work group concluded that cholesterol management for the reduction of cardiovascular risk was too important to remove completely from the composite measure aimed at reducing modifiable risk factors. The group is proposing to move forward with a redesign of this component in a thoughtful, staged approach. After several meetings in 2014 and a thorough discussion of new guidelines, evidence, safety and patient preference, the workgroup completed the cholesterol component redesign which was reviewed and approved by the MNCM Measurement and Reporting Committee in October 2014. http://mncm.org/cholesterol-component-redesign-approved-by-marc/

Members of the measure development workgroup included:

Beth Averbeck, MD Chair Internal Med & MNCM Board,HealthPartners
Mark Nyman, MD Internal Med & MARC member,Mayo Clinic & Health System
Victor Montori, MD Endocrinology, Mayo Clinic
JoAnn Sperl-Hillen, MD Internal Med, HealthPartners
Courtney Baechler, MD Cardiologist, Allina Penny George Institute
Jonathon W. Godsall, MD Endocrinology, Allina Medical Group
Christopher Restad, DO Family Medicine, Health East
Rebecca Moxness, MD Endocrinology, Park Nicollet
Thomas Knickelbine, MD Cardiologist, Minneapolis Heart Institute
Woubeshet Ayenew, MD Cardiologist, Hennepin County Med Center
Terry Murray, RN Data Analyst, Allina Medical Group
Jeanine Rosner, RN QI, Park Nicollet
Monica Simmer Health Plan, Metropolitan Health Plan
Pam York State Agency, MDH/ SQRMS
Kris Soegaard Consumer/ Empl/ MARC Member, MN Health Action Group
Collette Pitzen Facilitator/ Measure Dev MNCM

Members of the 2014 Measurement and Reporting Committee Included:
Tim Hernandez, MD Co-Chair/ Family Medicine, Medium Metro Medical Group
Howard Epstein, MD Co-Chair/ Hospitalist, Health Plan (PreferredOne)

Bill Nersesian, MD	Pediatrics, Large Metro Medical Group
Dan Walczak	Health Economics/analytics, Health Plan (Ucare)
Larry Lee, MD	Health Plan (Blue Cross Blue Shield MN)
Ann Robinow	Consumer/Health Policy Consultant
David Satin, MD	Family Medicine/ Researcher, Large Metro Medical Group
Laura Saliterman, MD	Pediatrics/Measurement/QI, Large Non-Metro Medical Group
Mark Nyman, MD	Internal Medicine, Large Non-Metro Medical Group
Caryn McGeary	Nursing/Measurement/QI, Small Non-Metro Clinics
Bruce Penner	Nursing/Measurement/QI, Medium Non-Metro Medical Group
Rahshana Price-Isuk, MD	Family Medicine/ Measurement/QI Safety Net Clinic
Chris Norton	Consumer (Retired teacher)
Mark Sonneborn	Hospital (MN Hospital Association)
Jeff Rank, MD	Gastroenterology, Single Specialty Group
Darin Smith	HEDIS Measurement/QI/Analyst, Health Plan (Medica)
Kris Soegaard	Purchasers(Buyers Health Care Action Group)
Allan Ross, MD	Family Medicine, Small Non-Metro Clinic
Matt Flory	Consumer (American Cancer Society)
Robert Lloyd	Purchaser, QI (MN Department of Human Services)
Sue Knudson	Informatics, Measurement/QI(HealthPartners)
Stefan Gildemeister	Health Policy, Data Analysis, QI(MN Department of Health)
David Homans, MD	Cardiology, Large Metro Medical Group/Hospitals

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 **Year the measure was first released:** 2002

Ad.3 **Month and Year of most recent revision:** 10, 2015

Ad.4 **What is your frequency for review/update of this measure?** Annual review

Ad.5 **When is the next scheduled review/update for this measure?** 10, 2016

Ad.6 **Copyright statement:**

Ad.7 **Disclaimers:**

Ad.8 **Additional Information/Comments:**

[i] Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors; National Quality Forum, Aug 2014. NQF Website

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 0288
De.2. **Measure Title:** Fibrinolytic Therapy Received Within 30 Minutes of ED Arrival
Co.1.1. **Measure Steward:** Centers for Medicare & Medicaid Services
De.3. **Brief Description of Measure:** This measure calculates the percentage of Emergency Department (ED) acute myocardial infarction (AMI) patients with ST-segment elevation on the electrocardiogram (ECG) closest to arrival time receiving fibrinolytic therapy during the ED stay and having a time from ED arrival to fibrinolysis of 30 minutes or less. The measure is calculated using chart-abstracted data, on a rolling, quarterly basis and is publicly reported, in aggregate, for one calendar year. The measure has been publicly reported, annually, by CMS as a component of its Hospital Outpatient Quality Reporting (HOQR) Program since 2012.
1b.1. **Developer Rationale:** Studies have shown that delays in the treatment of AMI leads to increased risk of in-hospital mortality and morbidity, with nearly 2 lives per 1,000 patients lost per hour of delay in treatment (Fibrinolytic Therapy Trialists' Collaborative Group, 1994). The total ischemic time—the time from onset of symptoms of ST-segment myocardial infarction (STEMI) to initiation of reperfusion therapy—is the principal determinant of patient health outcomes, so timely care is essential to reduce morbidity and mortality among STEMI patients. In situations in which it is deemed unlikely or impossible that a patient will receive primary PCI, the preferred treatment option, within the guideline-recommended time window of 120 minutes from first medical contact, fibrinolytic therapy is the recommended treatment (O'Gara, 2013). When fibrinolytic therapy is indicated or chosen as the appropriate reperfusion therapy in a patient presenting with a STEMI, it should be administered as rapidly as possible to reduce both mortality and morbidity rates, with a recommended time from hospital arrival to administration of 30 minutes (O'Gara, 2013). Accordingly, NQF #0228 assesses whether fibrinolytic therapy is administered within 30 minutes of ED arrival for patients with STEMI in order to ensure that all STEMI patients, regardless of their immediate access to a PCI-capable hospital, receive high-quality, time-sensitive care.

REFERENCES:

1) Fibrinolytic Therapy Trialists´ (FTT) Collaborative Group. Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomized trials of more than 1000 patients. Lancet. 1994; 343:311-22.

2) O´Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA, Granger CB, Krumholz HM, Linderbaum JA, Morrow DA, Newby LK, Ornato JP, Ou N, Radford MJ, Tamis-Holland JE, Tommaso CL, Tracy CM, Woo YJ, Zhao DX. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013 Jan 29;61(4):e78-140. Guideline available at: http://content.onlinejacc.org/article.aspx?articleid=1486115

S.4. **Numerator Statement:** The number of ED AMI patients whose time from ED arrival to fibrinolysis is 30 minutes or less.
S.7. **Denominator Statement:** The number of ED AMI patients with ST-segment elevation on ECG who received fibrinolytic therapy.
S.10. **Denominator Exclusions:** Patients are excluded who are less than 18 years of age. Additionally, patients who are not administered fibrinolytic therapy within 30 minutes AND had a Reason for Delay in Fibrinolytic Therapy, as defined in the Data Dictionary, are also excluded.

De.1. **Measure Type:** Process
S.23. **Data Source:** Administrative claims, Electronic Clinical Data : Electronic Health Record, Paper Medical Records
S.26. **Level of Analysis:** Facility, Population : National

IF Endorsement Maintenance – Original Endorsement Date: Nov 15, 2007 **Most Recent Endorsement Date:** Jan 18, 2012

Maintenance of Endorsement -- Preliminary Analysis

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

1a. Evidence. The evidence requirements for a *process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Summary of prior review in 2012:

- The steps between the measure focus and the health outcome include rapid assessment of patients experiencing a STEMI, determining the best reperfusion strategy, administer fibrinolytic therapy (if indicated/chosen) to STEMI patients as rapidly as possible (recommended time from hospital arrival to administration of 30 minutes)→reduction in mortality and morbidity.
- The developer provided one clinical guideline from the 2004 ACC/AHA Guidelines for the Management of Patients with ST-Elevation Myocardial Infarction:
  - The medical system goal is to facilitate rapid recognition and treatment of patients with STEMI such that door-to-needle (or medical contact–to-needle) time for initiation of fibrinolytic therapy can be achieved within 30 minutes. **Level of Evidence: A** (Level A: randomized controlled trial/meta-analysis)

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- The developer provided an additional clinical guideline from the 2013 ACCF/AHA Guideline for the Management of ST-Elevation Myocardial Infarction that provides three recommendations for fibrinolytic therapy when there is an anticipated delay to performing primary PCI within 120 minutes of first medical contact (FMC):
  1. In the absence of contraindications, fibrinolytic therapy should be administered to patients with STEMI at non-PCI-capable hospitals when the anticipated FMC-to-device time at a PCI-capable hospital exceeds 120 minutes because of unavoidable delays. **Class I, Level of Evidence: B**
  2. When fibrinolytic therapy is indicated or chosen as the primary reperfusion strategy, it should be administered within 30 minutes of hospital arrival. **Class I, Level of Evidence B**
  3. In the absence of contraindications, fibrinolytic therapy should be given to patients with STEMI and onset of ischemic symptoms within the previous 12 hours when it is anticipated that primary PCI cannot be performed within 120 minutes of FMC. **Class I, Level of Evidence: A**
- The developer provided a systematic review of the body of the evidence supporting the management of patients with ST-elevation myocardial infarction, specifically with reperfusion therapy. The details of the Quality, Quantity, and Consistency of the evidence provided was associated with the guideline (multiple randomized control trials).
- The developer identified three new studies that were published since the systematic review of the body of evidence (1986-2012).

Exception to evidence: N/A

Guidance from the Evidence Algorithm: Process measure/Systematic review (Box 3)→ QQC associated w/guideline presented (Box 4)→Systematic review concluded: Quantity: High; Quality: High; Consistency: High (Box 5a) →High

Questions for the Committee:

- *Is the Committee willing to accept the prior evaluation? The updated guideline supports the measure focus and has a strong level of evidence.*

Preliminary rating for evidence: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities
Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- The developer provided the following facility-level performance rates from Hospital Compare from the April 2010 – March 2015 data collection period:

	2010-11	2011-12	2012-13	2013-14	2014-15	2010-15 % Point Change
Facilities	121	109	103	79	76	-
Denominator	1,257	1,691	1,552	1,260	1,221	-
Minimum	9.0	17.0	9.0	9.0	9.0	-
1st PCTL	20.0	18.0	23.0	9.0	9.0	-11.0
5th PCTL	27.0	38.0	29.0	36.0	33.0	6.0
10th PCTL	38.0	42.0	40.0	39.0	42.0	4.0
25th PCTL	50.0	58.0	53.0	56.0	58.0	8.0
Median	65.0	70.0	69.0	73.0	73.0	8.0
75th PCTL	80.0	82.0	78.0	86.0	84.5	4.5
90th PCTL	91.0	91.0	91.0	95.0	93.0	2.0
95th PCTL	92.0	93.0	93.0	100.0	100.0	8.0
99th PCTL	100.0	100.0	97.0	100.0	100.0	-
Maximum	100.0	100.0	100.0	100.0	100.0	-
Weighted Mean (SD)	65.2 (82.1)	68.5 (68.7)	66.4 (74.4)	68.8 (90.4)	71.3 (73.5)	6.1

- The developer noted that the performance scores represented facilities whose denominator counts met minimum case count requirements during the April 2010-March 2015 data collection periods but does not state the minimum case count.

Disparities:

- The developer analyzed the effect of patient and facility characteristics on the likelihood of patients receiving fibrinolytic therapy within 30 minutes of ED arrival. The analysis included 3,844 cases submitted to the CDW in 2014 that included a principal diagnosis of AMI, ST-segment elevation on ECG performed closest to ED arrival, and received fibrinolytic therapy.
- The results indicated that age, race, sex, and facility size were variables related to the timely delivery of fibrinolytic therapy:
  - Patients aged 40-80 were significantly more likely than patients 18-30 to receive fibrinolytic therapy w/in 30 minutes of ED arrival.

- African-American patients were significantly less likely than White patients to receive fibrinolytic therapy w/in 30 minutes of ED arrival.
- Hispanic patients were less likely than non-Hispanic patients to receive fibrinolytic therapy w/in 30 minutes of ED arrival.
- Female patients were less likely than male patients to receive fibrinolytic therapy w/in 30 minutes of ED arrival.
- Patients treated in facilities with 101-500 beds were significantly more likely to receive fibrinolytic therapy w/in 30 minutes of ED arrival than patients treated in facilities with fewer than 50 beds.

Questions for the Committee:

- *Does the data presented demonstrate that there continues to be a quality problem and variation with STEMI patients receiving timely fibrinolytic therapy upon arriving in the ED? Is there an opportunity for improvement?*
- *Is a national performance measure still warranted?*
- *Are aware of evidence that additional disparities exist in this area of healthcare?*

Preliminary rating for opportunity for improvement: ☒ High ☐ Moderate ☐ Low ☐ Insufficient

Committee pre-evaluation comments

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence to Support Measure Focus

Comments: **There is strong evidence that rapid fibrinolytic therapy improves outcomes in patients with STEMI

**Evidence backing this process is strong and applies directly.

Process is tightly linked to the desired outcome.

I know of no new evidence not included in summary.

**The evidence supports the rationale that rapid identification of STEMI, determining the best reperfusion plan and administering lytics when appropriate will lead to a reduction in morbidity and mortality.

**A preponderance of evidence suggests that IN AN ERA WHERE TT (thrombolytic therapy) WAS FIRST LINE RX FOR STEMI, TT administered within 30 minutes was associated with modestly improved survival vs. 30-60 minutes. HOWEVER, this calculus might be substantially different for a patient presenting to an institution where PCI (in-house or via transfer) is sometimes available. If PCI is preferred then there is a legitimate concern that patients might be better served by a policy which does not promote immediate reflexive action. Also, dichotomizing the outcome as within X minutes or not within X minutes has always been at odds with clinical evidence or statistical theory or practice. Where X=30, it might be a particularly bad idea today.

**There is grade A evidence to support this process measure.

1b. Performance Gap

Comments: **There is a performance gap and there are age, sex, race and ethnicity disparities.

**Data provided does show a continued gap in care. The ultimate goal of STEMI process of care is to provide timely reperfusion--whether using PCI or lytic. 94% of patients with STEMI in US are treated with primary PCI. This measure addresses the other 6%. My concern with this small proportion is that it represents a mix of hospitals where lytic is the primary reperfusion strategy and others where it is used only in unusual circumstances. At the hospitals where it is the primary strategy, then this is a useful measure. At the ones where it is a secondary strategy used only under unusual circumstances, then it will not accurately reflect performance of their STEMI system of care. This is particularly important since it is publicly reported.

An ideal metric would combine the 2 STEMI reperfusion metrics into one.

**There are important disparities in care. The measure shows that females, African Americans, Hispanics,young patients and patients in smaller hospitals do not receive lytics as often.

**Data is provided for just 1,211 STEMI patients for 2015, treated at 76 facilities. This is <<1% of all STEMIs, <7% of patients receiving TT for STEMI, and <3% of facilities treating STEMI patients. This could mean there is room for improvement for these 1,211 patients, but it might also suggest that the 30 minute goal is not appropriate. Statements regarding effect of race, sex, age do suggest that rapid dx of STEMI is measurably worse for nonwhites, females, and age<30. Magnitude of the problem cannot be determined from available information. The use of 30 minutes as a dichotomous outcome here is particularly problematic in terms of determining whether there is a clinically-relevant difference.

Stats on page 21 are misleading. 546,352 AMI diagnoses overstates STEMIs by several-fold. ""Minority of U.S. hospitals ... capable

of PCI ..."" does not consider STEMI systems of care employed by EMS or informal systems of transfer to PCI care. Proportion of recent STEMIs ultimately treated with TT can be rather easily estimated (e.g. Medicare claims, HCUP), if not already published, and might help illuminate the issue. It would be interested to look at this by facility. "
**The measure developer describes disparities in the adoption of the elements of this measure in treating people with STEMI But again, is it the disparities or other factors contributing to these differences in use.
1c. High Priority (previously referred to as High Impact)
Comments: **n/a
**n/a
**not applicable
**N/A
**DNA

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

2a1. Reliability Specifications
Maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented.

Data source(s): Administrative claims, Electronic Health Record, Paper Medical Records

Specifications:

- This is a facility-level measure.
- The numerator includes patients aged 18 or older who have ST-elevation on the ECG closest to ED arrival and who receive fibrinolytic therapy within 30 minutes or less of ED arrival.
- The denominator includes patients who have ST-elevation on the ECG closest to ED arrival and who receive fibrinolytic therapy.
    - Patients who are discharged or transferred to a short-term general hospital for inpatient care or to a Federal healthcare facility are included in the denominator.
- The denominator exclusions include:
    - Patients less than 18 years of age
    - Patients who are not administered fibrinolytic therapy within 30 minutes of ED arrival AND had documentation of a reason for delay in fibrinolytic therapy as defined in the data dictionary
- Discharge codes, evaluation and management (E/M) codes, and ICD-10 codes included in attachment NQF_0288_MeasureCodeSet.xlsx.
- The calculation algorithm is included.
- Instructions for sampling and guidance on minimum sample size are provided.
- Details for handling missing data are provided.
- An electronic data collection tool is available from vendors. Free data collection tools are also available on qualitynet.org.

Questions for the Committee :

- *Are all the data elements clearly defined? Are all appropriate codes included?*
- *Is the logic or calculation algorithm clear?*
- *Is it likely this measure is consistently implemented?*

2a2. Reliability Testing attachment
Maintenance measures – less emphasis if no new testing data provided

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers.

For maintenance measures, summarize the reliability testing from the prior review:

- At the time of the last review the measure was undergoing validation through the CMS Clinical Abstraction Center.

Describe any updates to testing: measure score level testing conducted – see below

SUMMARY OF TESTING

Reliability testing level ☒ **Measure score** ☐ **Data element** ☐ **Both**

Reliability testing performed with the data source and level of analysis indicated for this measure ☒ **Yes** ☐ **No**

Method(s) of reliability testing:

- The dataset included 76 facilities that submitted 1,221 cases (denominator after exclusions) to Hospital Compare from April 1, 2014-March 31, 2015. Of those, a total of **871** cases met the numerator.
- The developers used a beta-binomial model to assess the signal-to-noise ratio to test the reliability of the measure score. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in performance. The higher the reliability score, the greater is the confidence with which one can distinguish the performance of one facility from another. This is an appropriate test for measure score reliability. A reliability of 0.70 is generally considered a minimum threshold for reliability.

Results of reliability testing:

- The developer included a histogram of the distribution of the reliability scores for the facilities meeting the minimum case counts from April 2014-March 2015, though a value for minimum case count is not provided.
- Reliability scores for facilities [meeting unknown minimum case count requirement] ranged from 0.49 to 1.00 with a median reliability of 0.67. This result indicates that 67% of the measure results reflects performance and 33% reflects noise.

Guidance from the Reliability Algorithm: Precise specifications (Box 1)→Empirical reliability testing (Box 2) →Computed performance scores for measured entities (Box 4) →Appropriate method used (Box 5) →Moderate

Questions for the Committee:

- *Is the test sample adequate to generalize for continued widespread implementation?*
- *Does the signal-to-noise estimates (median and range) demonstrate acceptable reliability for an accountability measure?*
- *Is the reliability estimate related to the sample size?*
- *Do the results demonstrate sufficient reliability so that differences in facility performance can be identified?*

Preliminary rating for reliability: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

2b. Validity
Maintenance measures – less emphasis if no new testing data provided

2b1. Validity: Specifications

2b1. Validity Specifications. This section should determine if the measure specifications are consistent with the evidence.

Specifications consistent with evidence in 1a. ☒ **Yes** ☐ **Somewhat** ☐ **No**
Specification not completely consistent with evidence

Question for the Committee:

- *Are the specifications consistent with the evidence?*

2b2. Validity testing

2b2. Validity Testing should demonstrate the measure data elements are correct and/or the measure score

correctly reflects the quality of care provided, adequately identifying differences in quality.

For maintenance measures, summarize the validity testing from the prior review:

- At the time of the last review the measure was undergoing validation through the CMS Clinical Abstraction Center.

Describe any updates to validity testing: data element testing – see below

SUMMARY OF TESTING

Validity testing level ☐ **Measure score** ☒ **Data element testing against a gold standard** ☐ **Both**

Validity testing method:

- The dataset used for patient-level data element testing included a sample of 462 (denominator) cases submitted to the Clinical Data Warehouse (CDW) from 23 facilities from April 1, 2014-March 31, 2015. After exclusions, **28** cases remained in the denominator sample; 18 of those cases met all of the numerator requirements.
- Data element validity was conducted by assessing the level of agreement between facility abstraction and auditor (CDAC) abstraction (gold standard) and calculating a kappa statistic or Pearson correlation coefficient.
    - Kappa values range between 0 and 1.0 and are interpreted as degree of agreement beyond chance. By convention, a kappa > .70 is considered acceptable.
        - 0 No better than chance
        - 0.01-0.20 Slight
        - 0.21-0.40 Fair
        - 0.41-0.60 Moderate
        - 0.61-0.80 Substantial
        - 0.81-1.0 Almost perfect
    - P-values estimate the statistical significance associated with the test statistics. P-values of less than 0.001 suggest high levels of statistical significance and that the degree of agreement was not due to chance.

Validity testing results:

- The kappa and Pearson's correlation coefficient for 11 critical data elements ranged from 0.63 to 1.0.

Data Element	Kappa/Pearson's Statistic (p-value)
E/M Code [a]	1.00 (<0.001)
Discharge Code [a]	1.00 (<0.001)
Patient Age on Outpatient Encounter Date [c]	-
ICD-9-CM Principal Diagnosis Code [a]	1.00 (<0.001)
Initial ECG Interpretation [a]	0.63 (<0.001)
Fibrinolytic Administration [a]	0.93 (<0.001)
Fibrinolytic Administration Date [b]	1.00 (<0.001)
Fibrinolytic Administration Time [b]	1.00 (<0.001)
Arrival Time [b]	1.00 (<0.001)
Time to Fibrinolysis [c]	-
Reason for Delay in Fibrinolytic Therapy [a]	1.00 (0.01)

a. The test statistic to assess validity for this data element is a Kappa score.
b. The test statistic to assess validity for this data element is a Pearson's correlation.
c. This data element is a calculated value, not an abstracted value.

Questions for the Committee:

○ *Do the validity testing results demonstrate that the measure elements are correct?*

○ *Is the test sample adequate to generalize for continued widespread implementation?*

2b3-2b7. Threats to Validity

2b3. Exclusions:

- The developer examined overall frequencies and proportions of exclusions/exceptions for 3,024 facilities that submitted eligible cases in 2014. The sampled population included 64,826 patients, age 18 years or older, who presented with an acute myocardial infarction with ST-elevation on ECG closest to ED arrival:

Overall Occurrence and Distribution across Facilities for Measure Exclusions and Exceptions

Data Element	Denominator Exclusion or Numerator Exception?		Overall Occurrence		Distribution Across Facilities (%)		
	Denominator Exclusion	Numerator Exception	N	%	25th	50th	75th
Initial ECG Interpretation	X		38,573	59.5	44.1	60.0	76.7
Fibrinolytic Administration	X		21,654	33.4	10.9	31.6	50.0
Fibrinolytic Administration Date*		X	3	0.0	0.0	0.0	0.0
Fibrinolytic Administration Time*		X	18	0.0	0.0	0.0	0.0
Arrival Time*		X	1	0.0	0.0	0.0	0.0
Time to Fibrinolysis	X		37	0.1	0.0	0.0	0.0
Reason for Delay in Fibrinolysis	X		718	1.1	0.0	0.0	0.0
Total Denominator Exclusions	4	-	60,982	94.1	93.8	100.0	100.0
Total Numerator Exceptions	-	3	22	0.0	0.0	0.0	0.0
Total Removed from the Denominator or Numerator	7 exceptions and exclusions		61,004	94.1	93.8	100.0	100.0

*If these data elements are to "UTD", the case is not included in the measure numerator but remains in the effective sample.

Questions for the Committee:

○ *Are the exclusions consistent with the evidence?*

○ *Does the large number of denominator exclusions (>94%)represent a significant threat to validity of the measure results?*

○ *Are any patients or patient groups inappropriately excluded from the measure?*

○ *Are the exclusions/exceptions of sufficient frequency and variation across providers to be needed (and outweigh the data collection burden)?*

2b4. Risk adjustment: **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

2b5. Meaningful difference *(can statistically significant and clinically/practically meaningful differences in performance measure scores can be identified)*:

- The developer provided the distribution of facility performance scores for 76 facilities from the April 2014-March 2015 data collection period.

Distribution of Facility Performance Scores								
Mean	Std. Dev.	Min.	10th Percent	Lower Quartile	Median	Upper Quartile	90th Percent	Max.

70.6	19.5	9.0	42.0	58.0	73.0	84.5	93.0	100.0

Question for the Committee:

- *Does this measure identify statistically significant and meaningful differences in performance among facilities?*

2b6. Comparability of data sources/methods:

- Measure is not specified for more than one data source; comparability of data sources is not needed.

2b7. Missing Data

- Cases with a value of "UTD" are included in the denominator but excepted from the numerator. The developer described the frequency and distribution of missing data in Section 2b3.3.

Guidance from Validity Algorithm: Specifications consistent with evidence (Box 1) →Potential threats to validity assessed; significant number of exclusions(Box 2) →Empirical validity testing (Box 3)→Patient-level data element validity testing conducted; only approximately 6% of cases remained in denominator sample after exclusions (Box 10) →Element data elements abstracted and compared against the gold standard and kappa/Pearson's correlation coefficient calculated (Box 11)→Moderate (highest eligible rating is MODERATE)

Preliminary rating for validity: ☐ **High** ☐ **Moderate** ☒ **Low** ☐ **Insufficient**

Rationale: Significantly large number of exclusions are a threat to validity of the measure; > 94% of patients are removed from the denominator due to the four exclusions.

Committee pre-evaluation comments

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2d)

2a1. & 2b1. Specifications

Comments: **"Data specifications were clearly defined.

**All are clearly defined and the logic is clear. However, need to update from ICD-9 basis to ICD-10 basis.

**The data elements are clearly defined, no concerns

**It is not clear to me why presence of ST elevation on EKG is part of the denominator specification. If IV TT is given for anything other than stroke symptoms, then it is because MD interpreted the EKG as showing ST elevation (there are some issues re: LBBB but none that changes this idea). Not surprising that ST elevation is frequently not documented. Not only are there a very large # of exclusions on this basis, the exclusion seems like an easy gaming opportunity (don't record EKG interpretation if TT administered late). If this measure is to be retained, it should be re-specified in terms of receiving TT for suspected STEMI, not EKG (note that the same logic doesn't apply to the PCI measure).

On page 38 there is reference to ""clear clinical need for timely TT administration."" To be clear, for TT, unlike PCI, the need is always timely, since there is a diminishing benefit over time which must outweigh a substantial risk of catastrophic intra-cranial bleed which does not decrease over time.

I cannot determine how exclusions are defined. Text refers to a data dictionary, but there is no link. It is unclear to me how the fibrinolytic therapy exclusion comes into play. If the numerator is defined in terms of STEMI patients who receive TT, who are the 20,000 patients excluded because they don't receive TT?

It is not clear to me why code values 197.XX (""during surgery"") are included in the code set."

**The validity testing is based on only 28 cases for a common problem. Thus, this testing itself is not reliable at all.

**The specification are consistent with the evidence

**As discussed in 1a, given current belief that primary PCI is better, it is not clear to me that the 30 minute door-to-needle goal is still appropriate-- it could have the unintended consequence of encouraging premature TT administration, where the more prudent course is to take extra minutes to determine whether timely PCI is available. Given an intracranial bleed rate of >1% for TT, it would be easy to tip the balance in the wrong direction if there was an attempt to comply. The fact that <10% of patients who receive TT meet the 30 minute target suggests that the target might not be appropriate, achievable, or either, in this subpopulation.

2a2. Reliability Testing

Comments: **Reliability was tested with a beta-binomial. reliability was moderate.

**I don't understand how you get a measurement error for a binary measure. What is signal and what is noise here?
The biggest threat to validity and reliability is not the measure itself but the fact that it is measuring what is overwhelmingly a secondary process for STEMI care. (See my response to 1b above.)"
**Do the signal to noise estimates demonstrate acceptable reliability for an accountability measure?
**I am skeptical that reliability scores have a straightforward interpretation in the context of >97% of facilities excluded for "inadequate cases."
2b2. Validity Testing
Comments: **Validity testing performed for chart abstraction.
**The validity testing is based on only 28 cases for a common problem. Thus, this testing itself is not reliable at all.
**Large number of exclusions are a threat to validity
**Various validity issues are noted in the previous sections. Note that the measures of inter-rater reliability (page 34) do not touch upon primary threats to validity. I cannot interpret the face validity survey (apparently of five respondents). I am not sure how to interpret the table on page 37. Does row 1 mean that top 75th percentile facilities in terms of performance had 77% exclusions on basis of not having ECG interpretation? If so, this would seem consistent with my point about gaming (see 2a1 above). Again, what does any of this mean given tiny # of facilities represented. Are these the same facilities year after year?
2b3. Exclusions Analysis
2b4. Risk Adjustment/Stratification for Outcome or Resource Use Measures
2b5. Identification of Statistically Significant & Meaningful Differences In Performance
2b6. Comparability of Performance Scores When More Than One Set of Specifications
2b7. Missing Data Analysis and Minimizing Bias
Comments: **"The proportion of patients excluded is around 94%. Even so, I'm not sure why patients who do not get fibrinolytic therapy are excluded. It seems to me that time to therapy should be ""infinity" There was no risk adjustment.
**Exclusions are consistent with evidence and no groups are inappropriately excluded. Yes, the burdens of collecting the exclusions are significant but they are common and important enough that they should remain.
No risk adjustment.
The biggest threat to validity and reliability is not the measure itself but the fact that it is measuring what is overwhelmingly a secondary process for STEMI care. (See my response to 1b above.) "
**Large number of exclusions are a threat to validity
**Again, various threats to validity have already been mentioned. 94% exclusion rate leaving only 1,221 cases suggests that either the clinical paradigm is ""topped out"" (those delayed are delayed for good reason) or exclusions need to be changed, or likely both.

Criterion 3. Feasibility
Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- Some data elements are in defined fields in electronic sources.
- Two data elements, Initial ECG Interpretation and Reason for Delay in Fibrinolytic Therapy, rely on logic and inferences that abstractors have been trained to interpret, therefore potential for e-specification will require special attention.
- Data collection burden due to manual abstraction to collect the data.
- There are no fees, licensure, or other requirements necessary to use this measure.
- The developer stated that the majority of a five member expert panel agreed practical aspects of reporting this chart-abstracted measure do not place undue burden on facilities that collect the data though the costs of collecting the data and reporting this measure were not provided.

Questions for the Committee:

- *Are the required data elements routinely generated and used during care delivery?*
- *Are the required data elements available in electronic form, e.g., EHR or other electronic sources?*

Preliminary rating for feasibility: ☐ **High** ☐ **Moderate** ☒ **Low** ☐ **Insufficient**

Rationale: Chart abstraction with little likelihood of electronic specifications anytime soon.

Committee pre-evaluation comments
Criteria 3: Feasibility

3a. Byproduct of Care Processes

3b. Electronic Sources

3c. Data Collection Strategy

Comments: **"Some data elements are in defined fields in electronic sources.

- Two data elements, Initial ECG Interpretation and Reason for Delay in Fibrinolytic Therapy, rely on logic and inferences that abstractors have been trained to interpret, therefore potential for e-specification will require special attention.
- Data collection burden due to manual abstraction to collect the data.
- There are no fees, licensure, or other requirements necessary to use this measure.
- The developer stated that the majority of a five member expert panel agreed practical aspects of reporting this chart-abstracted measure do not place undue burden on facilities that collect the data though the costs of collecting the data and reporting this measure were not provided

In particular, there is little likelihood of electronic specification soon.

**All are routinely generated in clinical care. However, the ECG interpretation and reasons for exclusion currently require manual extraction. ECG interpretation could eventually lend itself to EHR incorporation and extraction. Assessment of exclusions for reasons for delay in thrombolytic administration will always require manual abstraction.

**Difficult to capture data elements correctly

**See previous comments regarding capturing EKG diagnosis of ST elevation. If measure is to be retained, note that many of the timestamps (arrival, drug ordering/delivery) are now collected automatically and should be incorporated into the measure.

Criterion 4: Usability and Use
Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact /improvement and unintended consequences

4. Usability and Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No**

Accountability program details:

- CMS HOQR Program: The HOQR Program is a pay for quality data reporting program implemented by CMS for outpatient hospital services. Hospital quality of care information gathered through the HOQR Program is publicly available on the Hospital Compare website.

Improvement results :

- The developer provided a summary of statistics of performance scores from the April 2010 – March 2015 data collection periods. The developer also stated that these cases reflect only a subset of the patients eligible for the measure. Depending upon the facility's total case count, the facility may report all cases or a sample of cases; thus, the number of patients receiving high-quality healthcare as performance on the measure improves may be larger than the number of cases captured by the measure.

Unexpected findings (positive or negative) during implementation:

- There has been wide variation in facility performance from the inception of public reporting through March 2015. While median performance scores have increased, there is an ongoing opportunity for improvement in facility performance. Depending on the facility's total case count, the facility may report all cases or a sample of cases; thus, the number of patients receiving high-quality healthcare as performance on the measure improves is larger than the number of cases captured by the measure.

Potential harms:

- The developer has not identified any unintended consequences.

***Questions for the Committee**:*

- *How can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Does the large number of exclusions limit the usability of this measure?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

Preliminary rating for usability and use: ☐ **High** ☐ **Moderate** ☒ **Low** ☐ **Insufficient**

Rationale: Limited usability and use due to the significantly high number of exclusions which excluded >94% of the patients.

Committee pre-evaluation comments

Criteria 4: Usability and Use

4a. Accountability and Transparency

4b. Improvement

4c. Unintended Consequences

Comments: **The results are publicly reported on Hospital Compare and currently in use in the CMS HOQR program. However, >94% of all patients were excluded from the analysis

**Publicly reporting this measure of questionable validity and reliability may commonly give a distorted view of a hospital's STEMI reperfusion care and therefore a distorted view of its quality.

An ideal metric would combine the 2 STEMI reperfusion metrics into one. This would be important and provide an accurate view of the quality of care provided to STEMI patients. "

**The large number of exclusions limit the usability of this measure.

**Again, as covered previously it is not clear to me that the benefits of this measure, as it is currently specified, outweigh the potential harms. There is little practical concern, however, because it is applied to almost no-one.

Criterion 5: Related and Competing Measures

Related or competing measures:

- 0163 : Primary PCI received within 90 minutes of hospital arrival
- 0290 : Median Time to Transfer to Another Facility for Acute Coronary Intervention

Harmonization:

- #0163 is included in the Hospital Inpatient Quality Reporting (HIQR) Program as an electronically specified clinical quality measure (eCQM) and focuses on the timely initiation of PCI for a patient who arrives at a PCI-capable hospital.
- #0290 focuses on the timely transfer of patients who require PCI.

Pre-meeting public and member comments

-

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0288

Measure Title: Fibrinolytic Therapy Received Within 30 Minutes of ED Arrival

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: 4/29/2016

Instructions

- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- Respond to all questions as instructed with answers immediately following the question. All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 10 pages (*incudes questions/instructions*; minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- Health outcome: [3] a rationale supports the relationship of the health outcome to processes or structures of care. Applies to patient-reported outcomes (PRO), including health-related quality of life/functional status, symptom/symptom burden, experience with care, health-related behavior.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- Efficiency: [6] evidence not required for the resource use component.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the U.S. Preventive Services Task Force (USPSTF) grading definitions and methods, or Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.

5. 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 multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☐ Health outcome: Click here to name the health outcome

☐ Patient-reported outcome (PRO): Click here to name the PRO

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors

☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☒ Process: This measure calculates the proportion of patients diagnosed with acute myocardial infarction (AMI) receiving fibrinolytic therapy within 30 minutes of hospital arrival.

☐ Structure: Click here to name the structure

☐ Other:

HEALTH OUTCOME/PRO PERFORMANCE MEASURE *If not a health outcome or PRO, skip to 1a.3*

1a.2. Briefly state or diagram the path between the health outcome (or PRO) and the healthcare structures, processes, interventions, or services that influence it.

This measure is not a health outcome/PRO performance measure.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process, intervention, or service (*i.e., influence on outcome/PRO*).

This measure is not a health outcome/PRO performance measure.

Note: For health outcome/PRO performance measures, no further information is required; however, you may provide evidence for any of the structures, processes, interventions, or service identified above.

INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

1a.3. Briefly state or diagram the path between structure, process, intermediate outcome, and health outcomes. Include all the steps between the measure focus and the health outcome.

Time to fibrinolytic therapy is a strong predictor of outcome in patients with an acute myocardial infarction. Nearly 2 lives per 1,000 patients are lost per hour of delay (Fibrinolytic Therapy Trialists´ Collaborative Group, 1994). National guidelines recommend that fibrinolytic therapy be given within 30 minutes of hospital arrival in patients with ST-segment elevation myocardial infarction (Antman, 2004).

According to the Centers for Disease Control and Prevention (CDC) and American Heart Association (AHA), approximately 735,000 Americans suffer an acute myocardial infarction (AMI) every year, of which 29-47 percent are estimated to be ST-elevation myocardial infarctions (STEMIs) (Mozzaffarian, 2015). When a patient who is experiencing a STEMI makes first medical contact, it is critical for providers to rapidly assess the situation and determine the best reperfusion strategy. The total ischemic time—the time from onset of symptoms of STEMI to initiation of reperfusion therapy—is the principal determinant of patient health outcomes, so timely care is essential in order to reduce morbidity and mortality among STEMI patients (O'Gara, 2013).

While recommendations in national guidelines indicate that primary percutaneous coronary intervention (PCI) is the preferred early intervention for STEMI patients, a minority of U.S. hospitals are capable of providing PCI and disproportionate expansion of PCI-capable facilities across the country over the last decade has created gaps in access to primary PCI (Langabeer, 2013; O'Gara, 2013; Ellis, 2004). In situations in which it is deemed unlikely or impossible that a patient will receive primary PCI within the guideline-recommended time window of 120 minutes from first medical contact, fibrinolytic therapy is the preferred treatment (O'Gara, 2013). For patients that present within the first one to

two hours after symptom onset, immediate fibrinolysis may even be the preferred treatment to primary PCI, even when inter-facility transfer times are short (O'Gara, 2013).

Studies have shown that delays in treatment lead to increased risk of in-hospital mortality and morbidity, with nearly 2 lives lost per 1,000, per hour of delay in treatment (Fibrinolytic Therapy Trialists' Collaborative Group, 1994); longer time to fibrinolytic therapy was also associated with an increased infarct size (Bhatia, 2015). Therefore, when fibrinolytic therapy is indicated or chosen as the preferred reperfusion therapy in a patient presenting with a STEMI, it should be administered as rapidly as possible to reduce both mortality and morbidity rates, with a recommended time from hospital arrival to administration of 30 minutes (O'Gara, 2013). Accordingly, NQF #0228 assesses whether fibrinolytic therapy is administered within 30 minutes of ED arrival for patients with STEMI to ensure that all STEMI patients, regardless of their immediate access to a PCI-capable hospital, receive high-quality, time-sensitive care.

REFERENCES

Bhatia, V., R. Sood, D.S. Dhiman, A. Tomar, S. Asotra, P.C. Negi, and P. Panda. "Predictors of Acute Myocardial Infarct Size in STEMI Patients Receiving Thrombolytic Therapy: A Delayed Contrast Enhanced Cardiac MRI Study." Indian Heart Journal, vol. 67, no. 2, 2015, pp. 122–127.

Ellis SG, Armstrong P, Betriu A, Brodie B, Herrmann H, Montalescot G, Neumann FJ, Smith JJ, Topol E; on behalf of the Facilitated Intervention with Enhanced Reperfusion Speed to Stop Events (FINESSE) Investigators. Facilitated percutaneous coronary intervention versus primary percutaneous coronary intervention: design and rationale of the Facilitated Intervention with Enhanced Reperfusion Speed to Stop Events (FINESSE) trial. Am Heart J. 2004 Apr; 147(4):E16.

Fibrinolytic Therapy Trialists' (FTT) Collaborative Group. Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomized trials of more than 1000 patients. Lancet. 1994; 343:311-22.

Langabeer JR, Henry TD, Kereiakes DJ, Dellifraine J, Emert J, Wang Z, Stuart L, King R, Segrest W, Moyer P, Jollis JG. Growth in percutaneous coronary intervention capacity relative to population and disease prevalence. J Am Heart Assoc. 2013 Oct 28; 2(6):e000370.

Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, de Ferranti S, Despres JP, Fullerton HJ, Howard VJ, et al.; on behalf of the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2015 update: a report from the American Heart Association. Circulation. 2015;131:e30–e33.

O'Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA, et al. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013 Jan 29;61(4):e78-140.

1a.3.1. What is the source of the systematic review of the body of evidence that supports the performance measure?

☒ Clinical Practice Guideline recommendation – ***complete sections 1a.4, and 1a.7***

☐ US Preventive Services Task Force Recommendation – ***complete sections 1a.5 and 1a.7***

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*) – ***complete sections 1a.6 and 1a.7***

☒ Other – ***complete section 1a.8***

Please complete the sections indicated above for the source of evidence. You may skip the sections that do not apply.

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION

1a.4.1. Guideline citation (*including date*) and **URL for guideline** (*if available online*):

Antman EM, Anbe DT, Armstrong PW, Bates ER, Green LA, Hand M, Hochman JS, Krumholz HM, Kushner FG, Lamas GA, Mullany CJ, Ornato JP, Pearle DL, Sloan MA, Smith SC Jr. ACC/AHA guidelines for the management of patients with ST-elevation myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1999 Guidelines for the Management of Patients With Acute Myocardial Infarction). 2004.

The clinical practice guideline provided is based on its relevance to the measure. The guideline, released in 2013 by the American College of Cardiology Foundation (ACCF) and the American Heart Association (AHA), evaluates management of patients with STEMI. Citation for the guideline follows:

O'Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA, Granger CB, Krumholz HM, Linderbaum JA, Morrow DA, Newby LK, Ornato JP, Ou N, Radford MJ, Tamis-Holland JE, Tommaso CL, Tracy CM, Woo YJ, Zhao DX. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013 Jan 29;61(4):e78-140. Guideline available at: http://content.onlinejacc.org/article.aspx?articleid=1486115

1a.4.2. Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

"The medical system goal is to facilitate rapid recognition and treatment of patients with STEMI such that door-to-needle (or medical contact–to-needle) time for initiation of fibrinolytic therapy can be achieved within 30 minutes" Page 597

The ACCF/AHA guideline for the management of ST-elevation myocardial infarction provides recommendations for fibrinolytic therapy when there is an anticipated delay to performing primary PCI within 120 minutes of first medical contact (FMC). Three recommendations support the measure's clinical intent:

1. In the absence of contraindications, fibrinolytic therapy should be administered to patients with STEMI at non-PCI-capable hospitals when the anticipated FMC-to-device time at a PCI-capable hospital exceeds 120 minutes because of unavoidable delays (81, 87, 88). (Class I, Level of Evidence: B; pg. e86)
2. When fibrinolytic therapy is indicated or chosen as the primary reperfusion strategy, it should be administered within 30 minutes of hospital arrival (89-93) (Class I, Level of Evidence B; pg. e86).
3. In the absence of contraindications, fibrinolytic therapy should be given to patients with STEMI and onset of ischemic symptoms within the previous 12 hours when it is anticipated that primary PCI cannot be performed within 120 minutes of FMC (81, 306-311). (Class I, Level of Evidence: A; pg. e94)

1a.4.3. Grade assigned to the quoted recommendation with definition of the grade:

A ABC Scale

All relevant recommendations from the guideline received a Class I designation. The evidence (level of evidence *A*) strongly and unambiguously supports the recommendation to give fibrinolytic therapy to patients with STEMI when it is anticipated that primary PCI cannot be performed within 120 minutes of FMC. Additionally, there is a broad consensus in the medical community (level of evidence *B*) supporting the recommendations to administer fibrinolytic therapy to patients with STEMI when indicated within 30 minutes of hospital arrival. The ACCF/AHA Task Force on Practice Guidelines asserts that a recommendation with Level of Evidence *B* or *C* does not imply that the recommendation is weak, as many important clinical questions addressed in the guidelines do not lend themselves to clinical trials. Despite a limited pool of randomized control trials (RCTs), there may be clear clinical consensus that a particular test or therapy is useful or effective. Collectively, the evidence supports these recommendations, demonstrating consensus within the clinical community that fibrinolytic therapy is reasonable for patients with STEMI at a non-PCI-capable hospital when the anticipated FMC-to-device time at a PCI-capable hospital exceeds 120 minutes.

The following grading scale applies to recommendations from the guideline:

Recommendation 1: Class I: Benefit >>>Risk Procedure/Treatment **should** be performed/ administered.

The following evidence scales apply to recommendations from the guideline:

One class of recommendations: Class I

Class I: Recommendation that procedure or treatment is useful/ effective

Two levels of evidence: Level A and Level B

Level A: Data derived from multiple randomized clinical trials or meta-analyses

Level B: Data derived from a single randomized trial or nonrandomized studies

1a.4.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: If separate grades for the strength of the evidence, report them in section 1a.7.*)

Level A (randomized controlled trial/ meta-analysis):

High quality randomized controlled trial that considers all important outcomes. High-quality meta-analysis (quantitative systematic review) using comprehensive search strategies.

Level B (other evidence):

A well-designed, nonrandomized clinical trial. A nonquantitative systematic review with appropriate search strategies and well-substantiated conclusions. Includes lower quality randomized controlled trials, clinical cohort studies, and case-controlled studies with nonbiased selection of study participants and consistent findings. Other evidence, such as high-quality, historical, uncontrolled studies, or well-designed epidemiologic studies with compelling findings, is also included.

Level C (consensus/expert opinion):

Consensus viewpoint or expert opinion. Expert opinion is sometimes the best evidence available

Additional grading scale for the recommendations:

Class IIa: Benefit >>Risk additional studies with focused objectives needed. ***It is reasonable*** *to perform/administer treatment*
Class IIb: Benefit ≥ Risk additional studies with broad objectives needed: additional registry data would be helpful. Procedure/Treatment ***may be considered****.*
Class III No Benefit: Procedure/Test: not helpful, Treatment: no proven benefit
Class III Harm: Procedure/Test: excess cost w/o benefit or harmful, Treatment: harmful to patients

Additional evidence scales:

Level C: Very limited populations evaluated. Only consensus opinions of experts, case studies, or standard of care.

1a.4.5. Citation and URL for methodology for grading recommendations (*if different from 1a.4.1*)**:**

ACCF/AHA Task Force on Practice Guidelines. Manual for ACCF/AHA Guideline Writing Committees: Methodologies and Policies from the ACCF/AHA Task Force on Practice Guidelines. American College of Cardiology and American Heart Association. 2006. Available at:

http://assets.cardiosource.com/Methodology_Manual_for_ACC_AHA_Writing_Committees.pdf and http://my.americanheart.org/professional/StatementsGuidelines/PoliciesDevelopment/Development/Methodologies-and-Policies-from-the-ACCAHA-Task-Force-on-Practice-Guidelines_UCM_320470_Article.jsp.

1a.4.6. If guideline is evidence-based (rather than expert opinion), are the details of the quantity, quality, and consistency of the body of evidence available (e.g., evidence tables)?

☒ Yes → ***complete section 1a.7***

☐ No → ***report on another systematic review of the evidence in sections 1a.6 and 1a.7; if another review does not exist, provide what is known from the guideline review of evidence in 1a.7***

1a.5. UNITED STATES PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

1a.5.1. Recommendation citation (*including date*) and **URL for recommendation** (*if available online*):

This measure is not based on a United States Preventive Services Task Force Recommendation.

1a.5.2. Identify recommendation number and/or page number and **quote verbatim, the specific recommendation**.

This measure is not based on a United States Preventive Services Task Force Recommendation.

1a.5.3. Grade assigned to the quoted recommendation with definition of the grade:

This measure is not based on a United States Preventive Services Task Force Recommendation.

1a.5.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: the grading system for the evidence should be reported in section 1a.7.*)

This measure is not based on a United States Preventive Services Task Force Recommendation.

1a.5.5. Citation and URL for methodology for grading recommendations (*if different from 1a.5.1*)**:**

This measure is not based on a United States Preventive Services Task Force Recommendation.

Complete section 1a.7

1a.6. OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1. Citation (*including date*) and **URL** (*if available online*):

Guideline is evidenced based; details are provided in **Section 1a.7**.

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

Guideline is evidenced based; details are provided in **Section 1a.7**.

Complete section 1a.7

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

If more than one systematic review of the evidence is identified above, you may choose to summarize the one (or more) for which the best information is available to provide a summary of the quantity, quality, and consistency of the body of evidence. Be sure to identify which review is the basis of the responses in this section and if more than one, provide a separate response for each review.

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

Methodologic Approach for the Systematic Review that Supports the Guideline:

Members of the writing committee were appointed by the ACCF/AHA Task Force and selected from the American College of Physicians, American College of Emergency Physicians, and Society for Cardiovascular Angiography and Interventions, representing various areas of medical expertise. Strict adherence to the Task Force Relationship with Industry (RWI) policy was maintained throughout the consensus process.

The focus of the guideline is the management of patients with ST-elevation myocardial infarction, particular emphasis has been placed in areas such as reperfusion therapy and organization of regional systems of care. Panel members extensively reviewed the relevant literature, focusing on publications through November 2010, with additional, selected references added through August 2012. Evidence supporting each guideline recommendation was weighted and ranked against the ACCF/AHA grading system. Recommendations have been developed using evidence-based methodologies created by the Task Force.

1a.7.2. Grade assigned for the quality of the quoted evidence with definition of the grade:

Grade for the evidence provided from the guideline can be found in **Section 1a.4.3.**

1a.7.3. Provide all other grades and associated definitions for strength of the evidence in the grading system.

Grade for the evidence provided from the guideline can be found in **Section 1a.4.3.**

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range:

The references listed in the body of evidence span from 1986 to 2012.

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

The guideline does not explicitly indicate the specific number or type of study designs included in the body of evidence; however, it does reference several randomized control trials and prospective studies. In addition, one of the recommendations is Level *A*, which is defined as data from multiple randomized clinical trials or meta analyses; two of the recommendations are Level *B*, which is defined as data derived from a single randomized trial or nonrandomized studies. These three recommendations on fibrinolytic use included 13 unique citations.

1a.7.6. What is the overall quality of evidence across studies in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

The guideline provides three Class I recommendations, indicating that the benefits clearly outweigh the risks and the recommendation can be applied to most patients in most circumstances. The one Level *A* recommendation is based on multiple RCTs with no important limitations or exceptionally strong evidence from observational studies, and further evidence is unlikely to change the confidence in the estimate of the the effect. The two Level *B* recommendations are based on nonrandomized or a single RCT; although randomized trials may not be available, there is a clear clinical consensus of the estimate of the effect.

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) across studies in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

Given the high costs associated with complications of STEMI and the subsuquent rehabilitation, the overall net benefit of timely fibrinolytics therapy is a reduction in cost and a reduction in morbidity.

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

The guideline does not provide details about potential harms associated with fibrinolytic therapy for patients with STEMI that were identified in the body of evidence.

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for each new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

In addition to the guideline cited above, a review of the clinical literature was conducted during the measure contractor's annual review of the literature for additional evidence and/or new studies that relate to the measure. Citations and summaries for four items included in this review can be found in **Section 1a.8.2**. Some of these four studies have been published since the period of guideline development. Results cited in these studies are consistent across studies and with the guidelines cited above.

1a.8 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.8.1 What process was used to identify the evidence?

In addition to the guideline cited above, a review of the clinical literature and related policy was conducted during the measure contractor's annual review of the literature for additional evidence and/or new studies that support the measure's intent. The measure contractor identified relevant peer-reviewed publications by searching the PubMed MEDLINE database from January 1, 2013 to September 1, 2015, limiting the inclusion of results to those published in the English language and those that had abstracts available in PubMed. The search initially identified 949 articles; a further review by the contractor's clinical and measure-development team pared down results to only those directly relevant to the measure. These results, as well as an additional review of evidence published prior to the structured review's time window, resulted in the inclusion of three articles, included in the body of evidence below. Citations and summaries for the three items included in this review can be found in **Section 1a.8.2**.

1a.8.2. Provide the citation and summary for each piece of evidence.

Fibrinolytic Therapy Trialists´ (FTT) Collaborative Group. Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomized trials of more than 1000 patients. Lancet. 1994; 343:311-22.

Krumholz HM, Anderson JL, Bachelder BL, Fesmire FM, Fihn SD, Foody JM, et al. ACC/AHA 2008 performance measures for adults with ST-elevation and non-ST-elevation myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures (Writing Committee to Develop Performance Measures for ST-Elevation and Non-ST-Elevation Myocardial Infarction). J Am Coll Cardiol. 2008;52:2046-99.

Bhatia, V., R. Sood, D.S. Dhiman, A. Tomar, S. Asotra, P.C. Negi, and P. Panda. "Predictors of Acute myocardial Infarct Size in STEMI Patients Receiving Thrombolytic Therapy: A Delayed Contrast Enhanced Cardiac MRI Study." Indian Heart Journal, vol. 67, no. 2, 2015, pp. 122–127.
Bhatia et al. conducted post-treatment assessment of 26 patients with STEMI receiving thrombolytic therapy in order to determine predictors of large acute myocardial infarct size. The group found through univariate analysis that a delayed time to thrombolysis was significantly correlated with a larger final infarct size, indicating the importance of early initiation of therapy to improve outcomes.

Russo, J.J., S.G. Goodman, W.J. Cantor, M.K. Tan, B. Borgundvaag, D. Fitchett, V. Dzavik, R.T. Yan, J.J. Graham, S.R. Mehta, and A.T. Yan. "Efficacy and Safety of a Routine Early Invasive Strategy in Relation to Time from Symptom Onset to Fibrinolysis (a Subgroup Analysis of TRANSFER-AMI)." American Journal of Cardiology, vol. 115, no. 8, 2015, pp. 1005–1012.
Russo et al. conducted a randomized control trial of 1,059 patients receiving fibrinolysis for STEMI, analyzing the efficacy and safety of early invasive strategy post-fibrinolysis versus standard therapy in relation to time from symptom onset to fibrinolysis. The group found that patients who received fibrinolytic therapy within two hours of symptom onset had lower Global Registry Acute Coronary Events risk scores than those who received fibrinolysis two or more hours after symptom onset; however, they concluded that there was no difference in the efficacy and safety of an early invasive strategy for STEMI patients in relation to the time from symptom onset to fibrinolysis.

Vora, A.N., D.N. Holmes, I. Rokos, M.T. Roe, C.B. Granger, W.J. French, E. Antman, T.D. Henry, L. Thomas, E.R. Bates, and T.Y. Wang. "Fibrinolysis Use Among Patients Requiring Interhospital Transfer for ST-Segment Elevation Myocardial Infarction Care: A Report from the U.S. National Cardiovascular Data Registry." Journal of the American Medical Association: Internal Medicine, vol. 175, no. 2, 2015, pp. 207–215.
Vora et al. conducted a retrospective analysis of 22,481 patients eligible for fibrinolysis or primary PCI that were transferred to a STEMI receiving center in the Acute Coronary Treatment and Intervention Outcomes Network-Get with the Guidelines database over a four year period. Their analysis found that neither the administration of primary PCI nor fibrinolytics is sufficiently meeting guideline-recommended time targets for reperfusion. The median door-to-needle time for patients receiving pre-transfer fibrinolysis was 34 minutes, and only 52.7% of eligible patients with a transfer

drive time exceeding 60 minutes receiving fibrinolysis. While patients receiving fibrinolysis prior to transfer had a higher bleeding risk than patients receiving primary PCI, Vora et al. found fibrinolytic therapy to be a viable reperfusion option when the anticipated time to primary PCI exceeded 120 minutes.

1. Evidence, Performance Gap, Priority – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all subcriteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form
NQF_0288_MeasureEvidenceForm.docx

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- disparities in care across population groups.

1b.1. Briefly explain the rationale for this measure *(e.g., the benefits or improvements in quality envisioned by use of this measure)*
Studies have shown that delays in the treatment of AMI leads to increased risk of in-hospital mortality and morbidity, with nearly 2 lives per 1,000 patients lost per hour of delay in treatment (Fibrinolytic Therapy Trialists' Collaborative Group, 1994). The total ischemic time—the time from onset of symptoms of ST-segment myocardial infarction (STEMI) to initiation of reperfusion therapy—is the principal determinant of patient health outcomes, so timely care is essential to reduce morbidity and mortality among STEMI patients. In situations in which it is deemed unlikely or impossible that a patient will receive primary PCI, the preferred treatment option, within the guideline-recommended time window of 120 minutes from first medical contact, fibrinolytic therapy is the recommended treatment (O'Gara, 2013). When fibrinolytic therapy is indicated or chosen as the appropriate reperfusion therapy in a patient presenting with a STEMI, it should be administered as rapidly as possible to reduce both mortality and morbidity rates, with a recommended time from hospital arrival to administration of 30 minutes (O'Gara, 2013). Accordingly, NQF #0228 assesses whether fibrinolytic therapy is administered within 30 minutes of ED arrival for patients with STEMI in order to ensure that all STEMI patients, regardless of their immediate access to a PCI-capable hospital, receive high-quality, time-sensitive care.

REFERENCES:

1) Fibrinolytic Therapy Trialists´ (FTT) Collaborative Group. Indications for fibrinolytic therapy in suspected acute myocardial infarction: collaborative overview of early mortality and major morbidity results from all randomized trials of more than 1000 patients. Lancet. 1994; 343:311-22.

2) O´Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA, Granger CB, Krumholz HM, Linderbaum JA, Morrow DA, Newby LK, Ornato JP, Ou N, Radford MJ, Tamis-Holland JE, Tommaso CL, Tracy CM, Woo YJ, Zhao DX. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013 Jan 29;61(4):e78-140. Guideline available at: http://content.onlinejacc.org/article.aspx?articleid=1486115

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for endorsement maintenance. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included). This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*
Analysis of facility-level data from Hospital Compare downloadable files indicates that there is variation in the administration of fibrinolytic therapy within 30 minutes of ED arrival for cases where patients have a principal diagnosis associated with AMI, have ST-segment elevation on the ECG performed closest to ED arrival, and received fibrinolytic therapy. During the April 2010-March 2011 data collection period, performance scores ranged from 9.0% to 100.0%, with a weighted mean of 65.2% (i.e., on average, 65.2% of STEMI patients received fibrinolytic therapy within 30 minutes of ED arrival). During the April 2014-March 2015 data collection period, performance scores ranged from 9.0% to 100.0%, with a weighted mean of 71.3%. From April 2010-March 2015, there was a 9.4% (6.1 percentage point) increase in the weighted mean of performance scores.

The data presented below represent performance scores and descriptive statistics for the facilities whose denominator counts met minimum case count requirements during the April 2010-March 2015 data collection periods.

	Data Collection Period		Percentage Point Change									
	April 2010-March 2011		April 2011-March 2012		April 2012-March 2013		April 2013-March 2014		April 2014-March 2015		April 2010-March 2015	
Facilities	121	109	103	79	76	-						
Minimum Value	9.0	17.0	9.0	9.0	9.0	-						
1st Percentile	20.0	18.0	23.0	9.0	9.0	-11.0						
5th Percentile	27.0	38.0	29.0	36.0	33.0	6.0						
10th Percentile	38.0	42.0	40.0	39.0	42.0	4.0						
25th Percentile	50.0	58.0	53.0	56.0	58.0	8.0						
Median	65.0	70.0	69.0	73.0	73.0	8.0						
75th Percentile	80.0	82.0	78.0	86.0	84.5	4.5						
90th Percentile	91.0	91.0	91.0	95.0	93.0	2.0						
95th Percentile	92.0	93.0	93.0	100.0	100.0	8.0						
99th Percentile	100.0	100.0	97.0	100.0	100.0	-						
Maximum Value	100.0	100.0	100.0	100.0	100.0	-						
Weighted Mean Performance (Standard Deviation)			65.2 (82.1)		68.5 (68.7)		66.4 (74.4)		68.8 (90.4)		71.3 (73.5)	6.1
Number of cases who received fibrinolytic therapy (Denominator)			1,257	1,691	1,552	1,260	1,221	-				

From the inception of public reporting through the March 2015 data collection periods; there has been wide variation in facility performance. The interquartile range of facility scores has been consistently wide over the years, with ranges between 24.0 and 30.0 percentage points. While median performance scores have increased, there is an ongoing opportunity for improvement in facility performance.

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

Data have been included in Section 1b.2; these data represent national performance over time, from April 2010-March 2015 data collection periods.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for endorsement maintenance. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*

The effect of patient and facility characteristics on the likelihood of each patient being administered fibrinolytic therapy within 30 minutes of ED arrival for cases where patients have a principal diagnosis associated with AMI, have ST-segment elevation on the ECG performed closest to ED arrival, and received fibrinolytic therapy was evaluated using 2014 data submitted to the Clinical Data Warehouse (CDW). The impact of patient and facility characteristics for the 3,844 cases that met these criteria was assessed using a logistic regression model.

Primary results from the regression were related to patient demographics. Cases where patients were in age groups 40 to 50 (OR= 3.80, p= 0.03), 50 to 60 (OR= 3.85, p= 0.03), 60 to 70 (OR= 3.44, p= 0.04), and 70 to 80 (OR= 3.10, p= 0.06) were significantly more likely than patients in age group 18 to 30 to receive fibrinolytic therapy within 30 minutes of ED arrival. African-American patients were significantly less likely than White patients to have received fibrinolytic therapy within 30 minutes of ED arrival (OR= 0.60, p= 0.001). In comparison to non-Hispanic patients, Hispanic patients were less likely to have received fibrinolytic therapy within 30 minutes of ED arrival (OR= 0.65, p= 0.03). Finally, female patients were less likely than male patients to have received fibrinolytic therapy within 30 minutes of ED arrival (OR= 0.77, p< 0.001).

The delivery of fibrinolytic therapy within 30 minutes of ED arrival for cases where patients have a principal diagnosis of AMI, have ST-segment elevation on the ECG performed closest to ED arrival, and received fibrinolytic therapy varied by the characteristics of the facility where the patient received care. When compared to patients treated in facilities with fewer than 50 beds (a proxy for facility size), patients treated in facilities with 101-250 beds (OR= 1.74, p= 0.002) and 251 to 500 beds (OR= 2.02, p= 0.017) were significantly more likely to have received fibrinolytic therapy within 30 minutes of ED arrival.

The regression model identified subpopulations of patients and facilities for which there are statistically significant differences in the

likelihood of each patient in the initial patient population—those who have a principal diagnosis of AMI, have ST-segment elevation on the ECG performed closest to ED arrival, and received fibrinolytic therapy –being administered fibrinolytic therapy within 30 minutes of ED arrival; however, these disparities do not indicate a need for risk adjustment. Adjusting for these differences would mask underlying differences in quality of care. As this is a process measure, there should be no difference in the standard of care for these patients; these statistically significant differences are likely driven by variation in provider practice. Consequently, risk adjustment or stratification is not necessary or appropriate for this measure.

1b.5. **If no or limited data on disparities from the measure as specified is reported in 1b4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations.**
Not applicable, as there are available data for disparities in care.

1c. High Priority (previously referred to as High Impact)

The measure addresses:

- a specific national health goal/priority identified by DHHS or the National Priorities Partnership convened by NQF; OR
- a demonstrated high-priority (high-impact) aspect of healthcare (e.g., affects large numbers of patients and/or has a substantial impact for a smaller population; leading cause of morbidity/mortality; high resource use (current and/or future); severity of illness; and severity of patient/societal consequences of poor quality).

1c.1. **Demonstrated high priority aspect of healthcare**
A leading cause of morbidity/mortality, Patient/societal consequences of poor quality, Severity of illness
1c.2. **If Other:**

1c.3. **Provide epidemiologic or resource use data that demonstrates the measure addresses a high priority aspect of healthcare. List citations in 1c.4.**
According to the Centers for Disease Control and Prevention (CDC) and American Heart Association (AHA), approximately 735,000 Americans experience an AMI every year, of which 29-47 percent are estimated to be STEMIs (Mozzaffarian, 2015). Time to fibrinolytic therapy is a strong predictor of morbidity and mortality outcomes in patients with a STEMI. Nearly 2 lives per 1,000 patients are lost per hour of delay (Fibrinolytic Therapy Trialists´ Collaborative Group, 1994). National guidelines recommend that fibrinolytic therapy be given within 30 minutes of hospital arrival in patients with STEMI (O'Gara, 2013).

To complement a review of the literature, national statistics on ED use from the Healthcare Cost and Utilization Project (HCUP) National Emergency Department Sample (NEDS) were extracted to estimate the total population of patients with a principal diagnosis associated with AMI. In 2013, there were an estimated 546,352 ED visits with a primary diagnosis of AMI. Most patients were ages 65-84 (42.3%), followed by patients ages 45-64 (37.2%), 85+ (14.2%), 18-44 (5.6%), and finally patients under 18 (0.7%). There were more male patients (61.1%) than female patients (39.9%).

Additionally, the Department of Health and Human Services (DHHS) includes AMI care in a number of national programs, including the Million Hearts Campaign and Healthy People 2020. Both of these initiatives galvanize existing efforts and new programs to improve cardiovascular health and quality of life through prevention, detection, and treatment of AMI and strokes (Frieden and Berwick 2011; DHHS 2014). Reducing door-to-fibrinolytic administration time through the continued reporting of NQF #0288 can help DHHS achieve its Healthy People 2020 objective of "increas[ing] the proportion of eligible patients with heart attacks who receive fibrinolytic therapy within 30 minutes of hospital arrival" (DHHS 2014).

Furthermore, recent literature supports existing evidence that only a minority of U.S. hospitals are capable of providing PCI, the preferred treatment for patients with STEMI. Over the past decade, disproportionate expansion of PCI-capable facilities across the country has created gaps in access to primary PCI (Langabeer, 2013; O'Gara, 2013; Ellis, 2004). As fibrinolytic therapy is the recommended treatment when it is deemed unlikely or impossible that a patient will receive primary PCI within the guideline-recommended time window of 120 minutes from first medical contact, it is important to continue reporting NQF #0288 in order to improve the quality of care for all individuals experiencing a STEMI.

The literature and utilization data demonstrate that AMI treatment remains a high priority aspect of healthcare because AMI is a leading cause of morbidity and mortality that affects a large number of individuals, with severe consequences resulting from poor quality of care.

1c.4. **Citations for data demonstrating high priority provided in 1a.3**
1) Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, de Ferranti S, Després J-P, Fullerton HJ, Howard VJ,

Huffman MD, Judd SE, Kissela BM, Lackland DT, Lichtman JH, Lisabeth LD, Liu S, Mackey RH, Matchar DB, McGuire DK, Mohler ER 3rd, Moy CS, Muntner P, Mussolino ME, Nasir K, Neumar RW, Nichol G, Palaniappan L, Pandey DK, Reeves MJ, Rodriguez CJ, Sorlie PD, Stein J, Towfighi A, Turan TN, Virani SS, Willey JZ, Woo D, Yeh RW, Turner MB; on behalf of the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2015 update: a report from the American Heart Association. Circulation. 2015;131:(4):e29-322.

2) O´Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA, Granger CB, Krumholz HM, Linderbaum JA, Morrow DA, Newby LK, Ornato JP, Ou N, Radford MJ, Tamis-Holland JE, Tommaso CL, Tracy CM, Woo YJ, Zhao DX. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013 Jan 29;61(4):e78-140. Guideline available at: http://content.onlinejacc.org/article.aspx?articleid=1486115

3) Frieden TR and Berwick D. The "Million Hearts" initiative – preventing heart attacks and strokes. N Engl J Med. 2011;365:e27.

4) DHHS. 2014. HDS-19.1 Increase the proportion of eligible patients with heart attacks who receive fibrinolytic therapy within 30 minutes of hospital arrival. http://www.healthypeople.gov/node/4579/data_details

5) Langabeer JR, Henry TD, Kereiakes DJ, Dellifraine J, Emert J, Wang Z, Stuart L, King R, Segrest W, Moyer P, Jollis JG. Growth in percutaneous coronary intervention capacity relative to population and disease prevalence. J Am Heart Assoc. 2013 Oct 28;2(6):e000370.

6) Ellis SG, Armstrong P, Betriu A, Brodie B, Herrmann H, Montalescot G, Neumann FJ, Smith JJ, Topol E; on behalf of the Facilitated Intervention with Enhanced Reperfusion Speed to Stop Events (FINESSE) Investigators. Facilitated percutaneous coronary intervention versus primary percutaneous coronary intervention: design and rationale of the Facilitated Intervention with Enhanced Reperfusion Speed to Stop Events (FINESSE) trial. Am Heart J. 2004 Apr;147(4):E16.

1c.5. If a PRO-PM (e.g. HRQoL/functional status, symptom/burden, experience with care, health-related behaviors), provide evidence that the target population values the measured PRO and finds it meaningful. *(Describe how and from whom their input was obtained.)*

This measure is not a PRO-PM measure.

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the subcriteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

Cardiovascular, Cardiovascular : Acute Myocardial Infarction

De.6. Cross Cutting Areas *(check all the areas that apply):*

S.1. Measure-specific Web Page *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*

https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier2&cid=1196289981244

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)

This is not an eMeasure **Attachment:**

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or*

csv file in the suggested format preferred - if not, contact staff)
Attachment **Attachment:** NQF_0288_MeasureCodeSet.xlsx

S.3. **For endorsement maintenance**, please briefly describe any changes to the measure specifications since last endorsement date and explain the reasons.
NQF #0288 was first endorsed by NQF in November 2007. Since its most recent endorsement in 2012, the measure specifications have been updated to reflect clinical changes in appropriate STEMI diagnosis or ED evaluation and management (E/M) codes; to address stakeholder feedback; and to harmonize with a similar measure in the Hospital Inpatient Quality Reporting (HIQR) program.

In 2012, the data elements were updated to provide clarification in abstraction and updates were made to selected references, including an updated version of the referenced guidelines. The Discharge Status data element was changed to Discharge Code. Multiple clarifications and modifications were made to the Initial ECG Interpretation data element to eliminate false inclusions and exclusions, and the list of negative qualifiers and modifiers was updated to align with a measure reported in the HIQR Program. The Reason for Delay in Fibrinolytic Therapy data element was also updated to align with the HIQR specifications.

In 2013, as part of the annual measure maintenance and review process, clarification was added to the Measure Information Form (MIF) regarding the use of retrospective and concurrent data. The cited guidelines for the management of STEMI were updated. Changes were made to the Arrival Time data element for all HOQR measures, and brand names were removed from the Reason for Delay in Fibrinolytic Therapy data element to avoid the appearance of endorsement.

In 2014, left bundle branch block (LBBB) was removed from the measure as an inclusionary ECG finding in order to align with updates made in the 2013 STEMI guidelines (O'Gara, 2013).

In 2015, all ICD-9-CM diagnosis codes were updated to corresponding ICD-10-CM diagnosis codes.

S.4. **Numerator Statement** *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome)*
IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
The number of ED AMI patients whose time from ED arrival to fibrinolysis is 30 minutes or less.

S.5. **Time Period for Data** *(What is the time period in which data will be aggregated for the measure, e.g., 12 mo, 3 years, look back to August for flu vaccination? Note if there are different time periods for the numerator and denominator.)*
Numerator: Fibrinolytic administration within a three-month period, aggregated on a rolling, quarterly basis for ED AMI patients whose time from ED arrival to fibrinolysis is 30 minutes or less.

Denominator: Fibrinolytic administration within a three-month period, aggregated on a rolling, quarterly basis for ED AMI patients with ST-segment elevation on ECG who received fibrinolytic therapy.

S.6. **Numerator Details** *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
The numerator is defined by six evaluation and management (E/M) codes and 18 ICD-10-CM diagnosis codes included in the value set for this measure; these detailed lists can be found in the Excel workbook provided for Section S2b.

The numerator includes patients age 18 or older who have ST-elevation on the ECG closest to ED arrival and who receive fibrinolytic therapy within 30 minutes or less of ED arrival. There are no numerator exceptions.

S.7. **Denominator Statement** *(Brief, narrative description of the target population being measured)*
The number of ED AMI patients with ST-segment elevation on ECG who received fibrinolytic therapy.

S.8. **Target Population Category** *(Check all the populations for which the measure is specified and tested if any):*
Populations at Risk, Senior Care

S.9. **Denominator Details** *(All information required to identify and calculate the target population/denominator such as definitions,*

specific data collection items/responses , code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)
The denominator is defined by six evaluation and management (E/M) codes and 18 ICD-10-CM diagnosis codes included in the value set for this measure; these detailed lists can be found in the Excel workbook provided for Section S2b.

The denominator includes patients who are discharged or transferred to a short-term general hospital for inpatient care or to a Federal healthcare facility, who have ST-segment elevation on the ECG performed closest to ED arrival, and who receive fibrinolytic therapy.

S.10. **Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*
Patients are excluded who are less than 18 years of age. Additionally, patients who are not administered fibrinolytic therapy within 30 minutes AND had a Reason for Delay in Fibrinolytic Therapy, as defined in the Data Dictionary, are also excluded.

S.11. **Denominator Exclusion Details** *(All information required to identify and calculate exclusions from the denominator such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
Cases are excluded for any patients that meet any of the following criteria:

- Patients less than 18 years of age.
- Patients who did not receive Fibrinolytic Administration within 30 minutes (Fibrinolytic Administration Date and Fibrinolytic Administration Time (in minutes) minus Outpatient Encounter Date and Arrival Time (in minutes) is greater than 30 minutes) AND had a Reason for Delay in Fibrinolytic Therapy, as defined in the Data Dictionary.

S.12. **Stratification Details/Variables** *(All information required to stratify the measure results including the stratification variables, definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b)*
Not applicable; this measure does not stratify its results.

S.13. **Risk Adjustment Type** (Select type. Provide specifications for risk stratification in S.12 and for statistical model in S.14-15)
No risk adjustment or risk stratification
If other:

S.14. **Identify the statistical risk model method and variables** *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development and testing should be addressed with measure testing under Scientific Acceptability)*
Not applicable; this measure does not risk adjust.

S.15. **Detailed risk model specifications** *(must be in attached data dictionary/code list Excel or csv file. Also indicate if available at measure-specific URL identified in S.1.)*
Note: Risk model details (including coefficients, equations, codes with descriptors, definitions), should be provided on a separate worksheet in the suggested format in the Excel or csv file with data dictionary/code lists at S.2b.

S.15a. **Detailed risk model specifications** (*if not provided in excel or csv file at S.2b*)
No risk model specifications are provided, as risk adjustment or stratification is not necessary for this measure.

S.16. **Type of score:**
Other (specify):
If other: Percentage

S.17. **Interpretation of Score** *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*
Better quality = Higher score

S.18. **Calculation Algorithm/Measure Logic** *(Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.)*
This measure calculates the percentage of ED AMI patients with ST-segment elevation on the ECG closest to arrival time receiving fibrinolytic therapy during the ED stay and having a time from ED arrival to fibrinolysis of 30 minutes or less. The patient population

is determined from two algorithms; the AMI Hospital Outpatient Population algorithm as well as the NQF #0288 measure-specific algorithm. The measure is calculated based on four consecutive quarters of hospital outpatient claims data, as follows:

1. Check E/M Code; if on Table 1.0 (in the Excel workbook provided for Section S2b), proceed
2. Check Discharge Code; include patients with discharge code of 4a or 4d
3. Calculate Patient Age (Outpatient Encounter Date - Birthdate)
4. Check Patient Age; if >= 18, proceed
5. Check ICD-10-CM Principal Diagnosis Code; if on Table 1.1 (in the Excel workbook provided for Section S2b), proceed to the measure-specific algorithm
6. Check Initial ECG Interpretation; if "Yes," proceed
7. Check Fibrinolytic Administration; if "Yes," proceed, record as the denominator
8. Check Fibrinolytic Administration Date; if a Non-Unable to Determine (UTD) value, proceed
9. Check Fibrinolytic Administration Time; if a Non-UTD value, proceed
10. Check Arrival Time; if a Non-UTD value, proceed
11. Calculate Time to Fibrinolysis (Fibrinolytic Administration Time minus Arrival Time)
12. Check Time to Fibrinolysis; if >= 0 min and <= 30 min, record as the numerator. If > 30 min and = 360 min, proceed
13. Check Reason for Delay in Fibrinolytic Therapy; if "Yes," patient is excluded from measure population. If "No," record in the denominator. Aggregate denominator and numerator counts by Medicare provider number
14. Measure = numerator counts / denominator counts [The value should be recorded as a percentage]

S.19. **Calculation Algorithm/Measure Logic Diagram URL or Attachment** *(You also may provide a diagram of the Calculation Algorithm/Measure Logic described above at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
Available at measure-specific web page URL identified in S.1

S.20. **Sampling** *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*
IF a PRO-PM, identify whether (and how) proxy responses are allowed.
Sampling is a process of selecting a representative part of a population in order to estimate the hospital's performance without collecting data for its entire population. Using a statistically valid sample, a hospital can measure its performance in an effective and efficient manner. Sampling is a particularly useful technique for performance measures that require primary data collection from a source such as the medical record. Sampling should not be used unless the hospital has a large number of cases in the outpatient population because a fairly large number of cases are needed to achieve a representative sample of the population. For the purpose of sampling outpatient department quality measures, the terms "sample," "effective sample," and "case" are defined below:
• The "sample" is the fraction of the population that is selected for further study.
• "Effective sample" refers to the part of the sample that makes it into the denominator of an outpatient measure set. This is defined as the sample for an outpatient measure set minus all the exclusions and contraindications for the outpatient measure set in the sample.
• A "case" refers to a single record (or an encounter) within the population. For example, during the first quarter a hospital may have 100 patients who had a principal diagnosis associated with the OP-1, 2, 3, 4, and 5 measures. The hospital's outpatient population would include 100 cases or 100 outpatient records for these measures during the first quarter.

To obtain statistically valid sample data, the sample size should be carefully determined, and the sample cases should be randomly selected in such a way that the individual cases in the population have an equal chance of being selected. Only when the sample data truly represent the whole population can the sample-based performance outpatient measure set data be meaningful and useful. Each hospital is ultimately responsible for adhering to the sampling requirements outlined in the specifications manual, available at Web page URL identified in Section S.1.

As a general rule/policy of CMS, hospitals are encouraged to submit as many cases as possible up to the entire population of cases if reasonably feasible. For example, if the raw data can be easily extracted from an existing electronic database or the abstraction burden is manageable, hospitals should consider submitting the entire population of cases that meet the initial selection criteria. Otherwise, a statistically valid sample can be selected.

S.21. **Survey/Patient-reported data** *(If measure is based on a survey, provide instructions for conducting the survey and guidance on minimum response rate.)*
IF a PRO-PM, specify calculation of response rates to be reported with performance measure results.
This measure does not use survey data.

S.22. **Missing data** (specify how missing data are handled, e.g., imputation, delete case.)
Required for Composites and PRO-PMs.
The measure does not make any adjustments for missing data. If data are missing, the case will proceed to Measure Category Assignment of X and will be rejected. While abstractors cannot submit missing data, they may submit a value of "UTD" for select data elements. Depending on the data element the case is then either excluded from the denominator or excepted from the numerator. Frequency and distribution of data with a value of "UTD" are reported in the attached Measure Testing Form.

S.23. **Data Source** *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*
If other, please describe in S.24.
Administrative claims, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

S.24. **Data Source or Collection Instrument** *(Identify the specific data source/data collection instrument e.g. name of database, clinical registry, collection instrument, etc.)*
IF a PRO-PM, identify the specific PROM(s); and standard methods, modes, and languages of administration.
An electronic data collection tool is made available from vendors or facilities can download the free CMS Abstraction & Reporting Tool (CART). Paper tools for manual abstraction, which are posted on www.QualityNet.org, are also available for the CART tool. These tools are posted on www.QualityNet.org.

S.25. **Data Source or Collection Instrument** *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
Available at measure-specific web page URL identified in S.1

S.26. **Level of Analysis** *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*
Facility, Population : National

S.27. **Care Setting** *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*
Hospital/Acute Care Facility
If other:

S.28. **COMPOSITE Performance Measure** - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*
Not applicable; this is not a composite measure.

2a. Reliability – See attached Measure Testing Submission Form
2b. Validity – See attached Measure Testing Submission Form
NQF_0288_MeasureTestingForm.docx

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b2-2b7)

Measure Number (*if previously endorsed*)**:** 0288
Measure Title: Fibrinolytic Therapy Received Within 30 Minutes of ED Arrival
Date of Submission: 4/29/2016
Type of Measure:

☐ Composite – ***STOP – use composite testing form***	☐ Outcome (*including PRO-PM*)
☐ Cost/resource	☒ Process
☐ Efficiency	☐ Structure

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- **For all measures, sections 1, 2a2, 2b2, 2b3, and 2b5 must be completed.**

- **For outcome and resource use measures**, section **2b4** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b6** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b2-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 20 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address sociodemographic variables and testing in this form refer to the release notes for version 6.6 of the Measure Testing Attachment.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **PRO-PMs and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b2. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **PRO-PMs and composite performance measures**, validity should be demonstrated for the computed performance score.

2b3. Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; [12]

AND

If patient preference (e.g., informed decision making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, 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). [13]

2b4. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and sociodemographic factors) that influence the measured outcome and are present at start of care; [14,15] and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b5. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** [16] **differences in performance**;

OR

there is evidence of overall less-than-optimal performance.

2b6. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b7. For **eMeasures, composites, and PRO-PMs** (or other measures susceptible to missing data), analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements 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 of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality.

12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.
13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions
15. 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 percent v. 75 percent) 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 less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE
Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (***must be consistent with data sources entered in S.23***)	**Measure Tested with Data From:**
☒ abstracted from paper record	☒ abstracted from paper record
☒ administrative claims	☒ administrative claims
☐ clinical database/registry	☐ clinical database/registry
☒ abstracted from electronic health record	☒ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

a) Datasets used to define the sample:

- The initial patient population is identified using chart-abstracted data for a sample of ED encounters with at least one of the following Current Procedural Terminology (CPT) codes for evaluation and management (E/M): 99281, 99282, 99283, 99284, 99285, or 99291. The initial patient population includes cases for patients 18 years and older, as of the date of the encounter, with a principal diagnosis associated with an acute myocardial infarction (AMI), identified by using any of the following International Classification of Diseases version 9 (ICD-9) codes: 410.00, 410.01, 410.10, 410.11, 410.20, 410.21, 410.30, 410.31, 410.40, 410.41, 410.50, 410.51, 410.60, 410.61, 410.70, 410.71, 410.80, 410.81, 410.90, or 410.91[1].

b) Datasets used to define the effective sample:

- The denominator is identified using chart-abstracted data for a sample of cases for patients included in the initial patient population.

c) Datasets used to identify denominator exclusions:

- Denominator exclusions are identified using chart-abstracted data of cases for patients included in the effective sample. Denominator exclusions capture cases for patients where any of the following conditions are met:
    - Cases where patients did not have ST-segment elevation on the ECG closest to ED arrival
    - Cases where patients did not receive fibrinolytic therapy
    - Cases where patients had a *Time to Fibrinolysis* value of less than zero minutes or greater than 360 minutes
    - Cases where patients did not receive fibrinolytic therapy within 30 minutes AND had a *Reason for Delay in Fibrinolytic Therapy*

d) Datasets used to capture the numerator:

- The numerator is identified using chart-abstracted data of cases for patients included in the denominator. The numerator includes cases for patients where the following conditions are met:
    - Cases where patients had ST-elevation on the ECG closest to ED arrival
    - Cases where patients receive fibrinolytic therapy within 30 minutes or less of ED arrival

e) Datasets used to identify numerator exceptions:

- Numerator exceptions are identified using chart-abstracted data of cases for patients included in the effective sample. Numerator exceptions include cases for patients when any of the following conditions are met:
    - *Fibrinolytic Administration Date* is equal to "UTD"
    - *Fibrinolytic Administration Time* is equal to "UTD"
    - *Arrival Time* is equal to "UTD"
    - Cases where patients did not receive fibrinolytic therapy within 30 minutes AND did not have a *Reason for Delay in Fibrinolytic Therapy*

1.3. What are the dates of the data used in testing? April 2010-March 2015

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of:	Measure Tested at Level of:

[1] As of October 1, 2015 the measure is calculated using ICD-10 codes. However, as these data will not be available in the CDW until 2016, testing was conducted on the most recent version of the measure using ICD-9 codes.

(*must be consistent with levels entered in item S.26*)	
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☒ hospital/facility/agency	☒ hospital/facility/agency
☐ health plan	☐ health plan
☒ other: national	☒ other: national

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

The number of measured entities (hospital emergency departments) varies by testing type; see **Section 1.7** for details.

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)
The number of patients varies by testing type; see **Section 1.7** for details.

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

Reliability Testing
Currently undergoing validation through the CMS Clinical Data Abstraction Center

Reliability Testing:
Data Source: Hospital Compare downloadable file [maintained by the Centers for Medicare & Medicaid Services (CMS)]
Dates: Denominator: April 1, 2014-March 31, 2015; Numerator: April 1, 2014-March 31, 2015; Exclusions: April 1, 2014-March 31, 2015; Exceptions: April 1, 2014-March 31, 2015
Number of Facilities: 76
Effective Sample (denominator after exclusions): 1,221
Numerator Cases: 871
Level of Analysis: Facility
Patient Characteristics: Not applicable

Validity Testing
Currently undergoing validation through the CMS Clinical Data Abstraction Center

Validity Testing – *Data Element Validity*
Data Source: Denominator: Clinical Data Warehouse (CDW); Numerator: CDW; Exclusions: CDW; Exceptions: CDW [maintained by CMS]
Dates: Denominator: April 1, 2014-March 31, 2015; Numerator: April 1, 2014-March 31, 2015; Exclusions: April 1, 2014-March 31, 2015; Exceptions: April 1, 2014-March 31, 2015
Number of Facilities: 23[2]
Sample (denominator): 462
Effective Sample (denominator after exclusions): 28
Numerator Cases: 18
Level of Analysis: Case

[2] The cases used for data element validity testing were abstracted from the CDW by CDAC and represent a sample of all CDW cases. The CDAC abstraction sample included 196 facilities; 23 facilities had cases for patients who were eligible for NQF #0288.

Sampled Patient Characteristics: Gender (% Male): 63.4; Mean Age (Years): 62.8 (St. Dev.: 14.6); Race (% Minority): 10.7

Validity Testing – *Face Validity*
Data Source: Structured qualitative survey completed by the stroke and acute myocardial infarction expert work group (EWG) members
Date Collected: January 2016-February 2016
Number of Responses: 5
Respondent Characteristics: Respondents were asked to self-identify as one or more of the following categories: clinician (5); healthcare administration (1); management (1)

Exclusions Analysis
N/A

Exclusions Analysis
Data Source: Denominator: CDW; Numerator: CDW; Exclusions: CDW; Exceptions: CDW [maintained by CMS]
Dates: Denominator: January 1, 2014-December 31, 2014; Numerator: January 1, 2014-December 31, 2014; Exclusions: January 1, 2014-December 31, 2014; Exceptions: January 1, 2014-December 31, 2014
Number of Facilities: 964[3]
Sample (denominator): 64,826
Effective Sample (denominator after exclusions): 3,844
Numerator Cases: 2,291
Level of Analysis: Case
Denominator Patient Characteristics: Gender (% Male): 72.7; Mean Age (Years): 61.1 (St. Dev.: 12.1); Race (% Minority): 6.9

Risk Adjustment Strategy
N/A

Risk Adjustment/Stratification
N/A- No risk adjustment or stratification was performed.

Identification of Meaningful Differences in Performance
N/A

Identification of Statistically Significant & Meaningful Differences in Performance
Data Source: Denominator: CDW; Numerator: CDW; Exclusions: CDW; Exceptions: CDW [maintained by CMS]
Dates: Denominator: April 1, 2014-March 31, 2015; Numerator: April 1, 2014-March 31, 2015; Exclusions: April 1, 2014-March 31, 2015; Exceptions: April 1, 2014-March 31, 2015
Number of Facilities: 76
Effective Sample (denominator after exclusions): 1,221
Numerator Cases: 871
Level of Analysis: Facility
Patient Characteristics: Not applicable

Comparability of Multiple Data Sources/Methods
N/A

Comparability of Performance Scores when more than one Set of Specifications
N/A- This measure only uses one set of specifications.

[3] The CDW sample included 3,024 facilities; 964 facilities had cases for patients who were eligible for NQF #0288.

Missing Data Analysis and Minimizing Bias
Data Source: Denominator: CDW; Numerator: CDW; Exclusions: CDW; Exceptions: CDW [maintained by CMS]
Dates: Denominator: January 1, 2014-December 31, 2014; Numerator: January 1, 2014-December 31, 2014; Exclusions: January 1, 2014-December 31, 2014; Exceptions: January 1, 2014-December 31, 2014
Number of Facilities: 964
Sample (denominator): 64,826
Effective Sample (denominator after exclusions): 3,844
Numerator Cases: 2,291
Level of Analysis: Case
Denominator Patient Characteristics: Gender (% Male): 72.7; Mean Age (Years): 61.1 (St. Dev.: 12.1); Race (% Minority): 6.9

1.8 What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example, patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).

We assessed patient-level SDS factors as part of the regression model reported in **Section 1b.4**, which provides an overview of disparities in care for patient sub-populations. We based this analysis on SDS variables included in the CDW data:

- Age
- Gender
- Race
- Ethnicity

While an analysis of SDS factors is important in understanding differences in care for patient sub-populations, this measure is a process measure that is neither risk-adjusted nor risk-stratified. We determined that risk adjustment and risk stratification were not appropriate based on the current evidence base and the measure construct. Additional information on this determination is provided in **Section 2b4.2**.

2a2. RELIABILITY TESTING
***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

N/A

☒ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)
☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

N/A

Performance score reliability was calculated in accordance with the methods discussed in *The Reliability of Provider Profiling: A Tutorial* (2009). This approach calculates the ability of the measure to distinguish true differences between the performances of different facilities. Specifically, the testing calculated the signal-to-noise ratio for each facility meeting the minimum case count, established by the measure calculation contractor, during the April 2014-March 2015 data collection period, with higher scores indicating greater reliability. The reliability score is estimated using a beta-

binomial model, which is appropriate for the reliability testing of pass/fail measures. The reliability score for each facility is a function of the facility's sample size and score on the measure, and the variance across facilities.

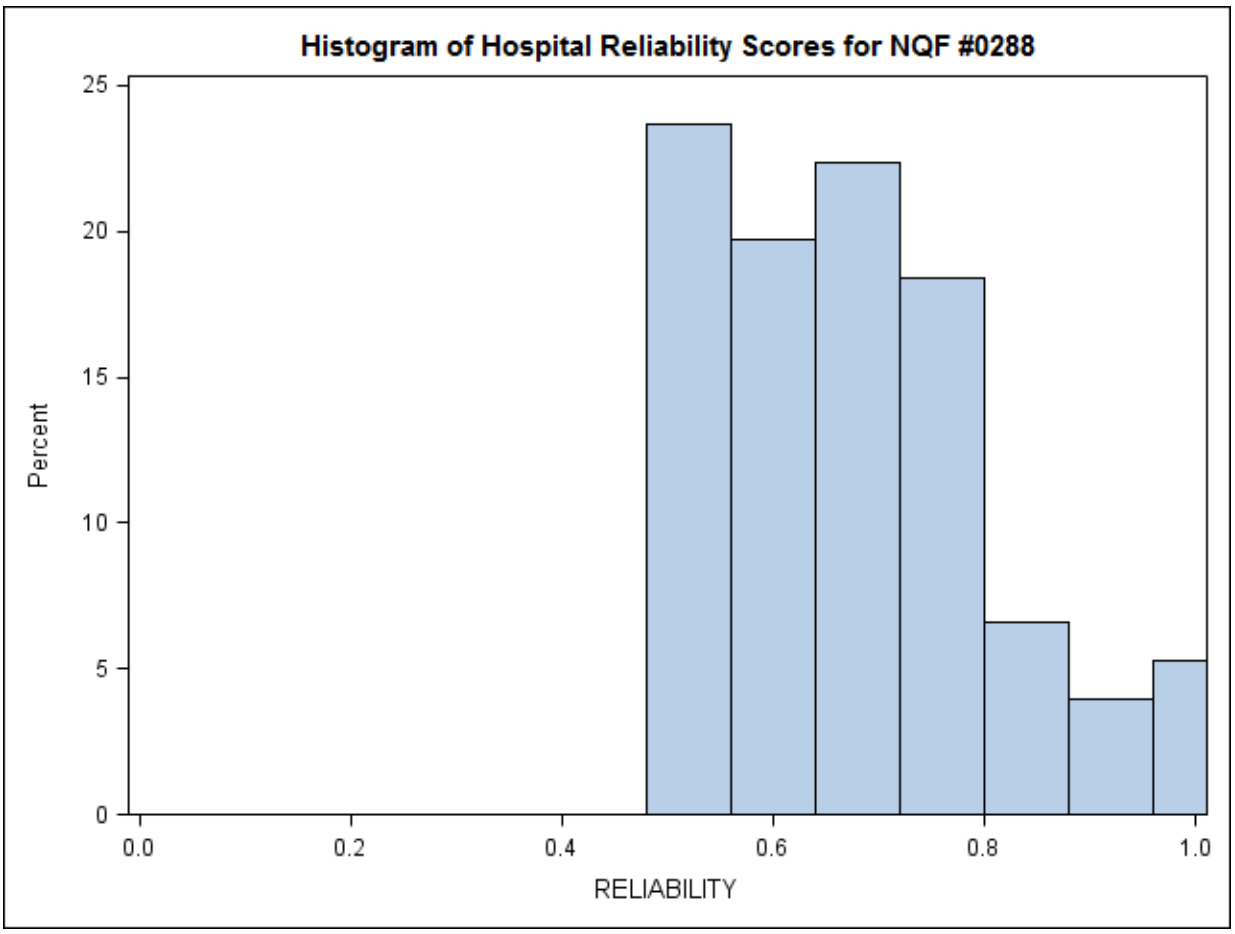
See **Section 2b2** for validity testing of data elements.

REFERENCE:
1) Adams JL. The reliability of provider profiling: a tutorial. Santa Monica, CA: RAND Corporation. 2009. Retrieved from http://www.rand.org/pubs/technical_reports/TR653.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

N/A

Figure 1 (below) is a histogram of the distribution of the performance score reliability for the facilities meeting the minimum case count requirements during the April 2014-March 2015 data collection period. Reliability scores ranged from 0.49 to 1.00, with a median reliability score of 0.67.



2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Calculated using a beta-binomial model, a median reliability score of 0.67 is indicative of moderate measure reliability. The results of this test indicate that the measure is able to identify true differences in performance between individual facilities.

2b2. VALIDITY TESTING

2b2.1. What level of validity testing was conducted? (*may be one or both levels*)

N/A

☒ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ **Performance measure score**

☐ **Empirical validity testing**

☒ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)

2b2.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

N/A

The validity of the measure was assessed using both qualitative and quantitative analyses. Data element validity of critical data elements was evaluated by calculating a kappa statistic (for categorical data elements) or Pearson correlation coefficient (for continuous data elements) assessing the level of agreement between facility abstraction and auditor (CDAC) abstraction. Face validity of the performance score was systematically assessed through survey of the EWG.

The test statistics measure interrater reliability and demonstrate the percent agreement between two sources for the same observation, after accounting for agreement by chance. For this test, CDAC is considered to be an authoritative source to which facility abstraction is compared. Test values range from 0.00 to 1.00, where a value of 0.00 indicates zero agreement between two sources and a value of 1.00 indicates complete agreement between two sources. To estimate the statistical significance associated with the test statistics, p-values can be calculated. P-values of less than 0.001 suggest very high levels of statistical significance, and suggest the results are not due to chance. For NQF measure #0288, kappa statistics and Pearson correlation coefficients were used to estimate the level of agreement between the facility's abstraction of critical data elements versus CDAC's abstraction of the critical data elements for the same sample of cases. The p-values associated with these estimates are also reported.

Landis & Koch, 1977 offer the following classification of kappa interpretation:

<0	Poor agreement
0.00–0.20	Slight agreement
0.21–0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–1.00	Almost perfect agreement

Similar interpretation is appropriate for Pearson correlation coefficients.

Face validity of the performance score was systematically assessed through survey of the EWG. Five EWG members participated in the data collection. Respondent perspectives include clinical, management, and healthcare administration. Prior to responding to questions related to measure-score face validity, EWG members were provided detailed measure specifications.

The following questions and statements related to measure-score face validity were posed to the EWG:

1. Patients who receive fibrinolytic therapy within 30 minutes of arrival to the ED can be accurately captured using chart-abstracted data.
2. The measure successfully assesses the timely administration of fibrinolytic therapy for AMI patients.

Responses to questions *1* and *2* in the measure-score face-validity section were collected using a five-point Likert scale: strongly agree, agree, undecided, disagree, strongly disagree, and do not know/not applicable.

REFERENCES:

2) Landis, J. & Koch, G. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159-174. 1977.

2b2.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

N/A

Results of critical data element validity testing indicate substantial to almost perfect agreement between the facility's abstraction of critical data elements versus CDAC's abstraction of the same sample of cases. Continued inclusion of a case in performance score calculation is dependent upon earlier data element values reported by the abstractor, and, therefore, it is possible that cases may have populated values for critical data elements after they have been excluded from the effective sample. For this reason, test statistics were only calculated for cases that remained in the effective sample at the corresponding step in the algorithm. For example, if a case had a value of "No" for *Initial ECG Interpretation* (thus excluding them from the effective sample) but also had populated values for later data elements, the case would not be considered in any calculations after *Initial ECG Interpretation*. The test statistic and p-value for each critical data element is provided in the table below, as well as the effective sample size used in the calculation.

Data Element	Test Statistic (p-value)	Effective Sample used in Kappa Calculation
E/M Code [a]	1.00 (<0.001)	462
Discharge Code [a]	1.00 (<0.001)	462
Patient Age on Outpatient Encounter Date [c]	-	-
ICD-9-CM Principal Diagnosis Code [a]	1.00 (<0.001)	462
Initial ECG Interpretation [a]	0.63 (<0.001)	462
Fibrinolytic Administration [a]	0.93 (<0.001)	141
Fibrinolytic Administration Date [b]	1.00 (<0.001)	28
Fibrinolytic Administration Time [b]	1.00 (<0.001)	28
Arrival Time [b]	1.00 (<0.001)	28
Time to Fibrinolysis [c]	-	-
Reason for Delay in Fibrinolytic Therapy [a]	1.00 (0.01)	25

a. The test statistic to assess validity for this data element is a Kappa score.
b. The test statistic to assess validity for this data element is a Pearson's correlation.
c. This data element is a calculated value, not an abstracted value.

Results of the face-validity assessment indicate that a diverse group of stakeholders support the validity of the measure. Results for each of the questions are provided below.

1. *Patients who receive fibrinolytic therapy within 30 minutes of arrival to the ED can be accurately captured using chart-abstracted data.*

Response Option	Response Percentage	Response Count
Strongly Agree	60.0%	3

Agree	40.0%	2
Undecided	0.0%	0
Disagree	0.0%	0
Strongly Disagree	0.0%	0
Do Not Know or Not Applicable	0.0%	0

2. The measure successfully assesses the timely administration of fibrinolytic therapy for AMI patients.

Response Option	Response Percentage	Response Count
Strongly Agree	40.0%	2
Agree	60.0%	3
Undecided	0.0%	0
Disagree	0.0%	0
Strongly Disagree	0.0%	0
Do Not Know or Not Applicable	0.0%	0

2b2.4. What is your interpretation of the results in terms of demonstrating validity? (i*.e., what do the results mean and what are the norms for the test conducted?*)

N/A

Results of the quantitative and qualitative analysis are positive and support the validity of the measure and its calculation. Based on the Landis and Koch classification scale, described in **Section 2b2.2**, there was substantial to almost perfect agreement between facility and auditor abstraction of data elements, with a majority of critical data elements (9 out of 11) falling in the almost perfect agreement range (0.81 to 1.00). Overall, test values ranged from 0.63-1.00 and were statistically significant (**Section 2b2.3**). This suggests strong validity for the measure, as currently specified.

The EWG, composed of five stakeholders representing healthcare administration, management, and clinicians with expertise in cardiology, neuro-radiology, emergency medicine, and emergency nursing, provided feedback on the face validity of NQF #0288 through an online survey. All members agreed or strongly agreed that patients who receive fibrinolytic therapy within 30 minutes of arrival to the ED can be accurately captured using chart-abstracted data. Similarly, they agreed or strongly agreed that NQF #0288 successfully assesses the timely administration of fibrinolytic therapy for AMI patients. The respondents generally support the face validity of NQF #0288.

2b3. EXCLUSIONS ANALYSIS

NA ☒ no exclusions — *skip to section 2b4*

NA ☐ no exclusions — *skip to section 2b4*

2b3.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

N/A

We tested measure exclusions and numerator exceptions to determine the prevalence of each exclusion and exception, by facility, and at an aggregate level. The analysis tested measure exclusions and numerator exceptions during the January 2014-December 2014 data collection period. Measure exclusions include all cases meeting one or more criteria listed in **Section 1.2c**, above. Numerator exceptions include cases meeting one or more criteria listed in **Section 1.2d**,

above. To supplement the empirical results, we systematically assessed the face validity of current exclusions and exceptions through survey of the EWG based on responses from five EWG members.

The face validity of exclusions was assessed, using the following statement:

1. The current measure specifications use chart-abstracted data to exclude patients from the denominator population based on the situations listed in the table below[2]. Please evaluate the appropriateness of the current exclusion criteria.

Responses to the question were collected using keep/remove response options.

2b3.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

N/A

We examined overall frequencies and proportions of cases excluded for each exclusion/exception criteria, among all sampled cases, for 3,024 facilities submitting eligible cases in 2014. The sampled population included 64,826 cases where a patient (age 18 years or older) presented with an acute myocardial infarction with ST-segment elevation on the ECG closest to ED arrival.

Overall Occurrence and Distribution across Facilities for Measure Exclusions and Exceptions

Data Element	Denominator Exclusion or Numerator Exception?		Overall Occurrence		Distribution Across Facilities (%)		
	Denominator Exclusion	Numerator Exception	N	%	25th	50th	75th
Initial ECG Interpretation	X		38,573	59.5	44.1	60.0	76.7
Fibrinolytic Administration	X		21,654	33.4	10.9	31.6	50.0
Fibrinolytic Administration Date		X	3	0.0	0.0	0.0	0.0
Fibrinolytic Administration Time		X	18	0.0	0.0	0.0	0.0
Arrival Time		X	1	0.0	0.0	0.0	0.0
Time to Fibrinolysis	X		37	0.1	0.0	0.0	0.0
Reason for Delay in Fibrinolysis	X		718	1.1	0.0	0.0	0.0
Total Denominator Exclusions	4	-	60,982	94.1	93.8	100.0	100.0

Data Element	Denominator Exclusion or Numerator Exception?		Overall Occurrence		Distribution Across Facilities (%)		
	Denominator Exclusion	Numerator Exception	N	%	25th	50th	75th
Total Numerator Exceptions	-	3	22	0.0	0.0	0.0	0.0
Total Removed from the Denominator or Numerator	7 exceptions and exclusions		61,004	94.1	93.8	100.0	100.0

The EWG provided feedback through an online survey on the appropriateness of the four categories of excluded populations: patients less than 18 years of age; patients without ST-elevation on the ECG performed closest to ED arrival; patients who did not receive fibrinolytic therapy; and patients who did not receive fibrinolytic therapy within 30 minutes AND who had a documented non-system (i.e., clinical or patient-centered) reason for delay in fibrinolytic therapy. The survey results indicate that a diverse group of stakeholders support the current exclusions for NQF #0288, strongly agree that no additional exclusions are necessary, and do not foresee significant challenges in capturing the exclusions described above using chart-abstracted data.

2b3.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

N/A

As seen in the table reported in **Section 2b3.2** above, the frequency of exclusions/exceptions varied substantially across facilities for two denominator exclusion criteria, while the frequency of the other five exclusions/exceptions were consistently low across facilities. Measure exclusion and numerator exception criteria are in alignment with clinical guidelines and also ensure that all cases included in the effective sample have sufficient denominator and numerator information to calculate the performance score. After identification of cases for patients 18 years and older with a principal diagnosis associated with acute myocardial infarction and ST-segment elevation on the ECG closest to ED arrival, measure exclusion and numerator exception criteria are applied:

- *Initial ECG Interpretation* is a denominator exclusion criterion. Cases for patients where *Initial ECG Interpretation* equals "No" are excluded from the effective sample. Overall, 59.5% of cases for patients included in the sample are excluded from the effective sample based on *Initial ECG Interpretation*. There is notable variability in the proportion of cases excluded based on *Initial ECG Interpretation* values across facilities, with an interquartile range of 44.1% to 76.7%. Only cases for patients with ST-segment elevation on the ECG closest to ED arrival are included in the effective sample because in such cases there is a clear clinical need for rapid administration of fibrinolytic therapy.

- *Fibrinolytic Administration* is a denominator exclusion criterion. Cases for patients where *Fibrinolytic Administration* equals "No" are excluded from the effective sample. Overall, 33.4% of cases included in the sample were excluded from the effective sample based on *Fibrinolytic Administration*. There is notable variability in the proportion of cases excluded based on *Fibrinolytic Administration* values across facilities, with an interquartile range from 10.9% to 50.0%. Only cases for patients who received fibrinolytic therapy are included in the effective sample because performance score is dependent upon the administration of fibrinolytics.

- *Fibrinolytic Administration Date* is a numerator exception criterion. If *Fibrinolytic Administration Date* is equal to "UTD," the case is not included in the measure numerator but remains in the effective sample. Overall, less than 0.1% of cases for patients included in the sample have a "UTD" value for *Fibrinolytic Administration Date*. While there is limited variability in the proportion of excepted cases based on *Fibrinolytic Administration Time* values across facilities, the exception remains important as a "UTD" value for this data element makes it impossible to determine whether the patient received timely fibrinolytic therapy.

- *Fibrinolytic Administration Time* is a numerator exception criterion. If *Fibrinolytic Administration Time* is equal to "UTD", the case is not included in the measure numerator but remains in the effective sample. Overall, less than 0.1% of cases for patients included in the sample have a "UTD" value for *Fibrinolytic Administration Time*. While there is limited variability in the proportion of excepted cases based on *Fibrinolytic Administration Time* values across facilities, the exception remains important as a "UTD" value for this data element makes it impossible to determine whether the patient received timely fibrinolytic therapy.

- *Arrival Time* is a numerator exception criterion. If *Arrival Time* is equal to "UTD", the case is not included in the measure numerator but remains in the effective sample. Overall, less than 0.1% of cases for patients included in the sample have a "UTD" value for *Arrival Time*. While there is limited variability in the proportion of excepted cases based on *Arrival Time* across facilities, the exception remains important as a "UTD" value for this data element makes it impossible to determine whether the patient received timely fibrinolytic therapy.

- *Time to Fibrinolysis* is a denominator exclusion criterion. Cases for patients where *Time to Fibrinolysis* is less than zero or more than 360 minutes are excluded from the effective sample. Overall, 0.1% of cases for patients included in the initial patient population are excluded from the effective sample based on *Time to Fibrinolysis*. There is limited variability in the proportion of excepted cases based on *Time to Fibrinolysis* values across facilities. Only cases for patients who receive fibrinolytics between zero and 360 minutes are included in the effective sample because a time less than zero minutes indicates poor documentation, and a time greater than 360 minutes may indicate that the need for fibrinolytics was not emergent.

- *Reason for Delay in Fibrinolysis* is a denominator exclusion criterion, conditional upon *Time to Fibrinolysis* greater than 30 minutes or less than or equal to 360 minutes. Cases for patients where *Reason for Delay in Fibrinolysis* is equal to "Yes" are excluded from the effective sample. Overall, 1.1% of cases for patients included in the sample are excluded from the effective sample based on *Reason for Delay in Fibrinolysis*. There is limited variability in the proportion of excepted cases across facilities. Only cases for patients who have a clinical or patient-centered reason for delay in fibrinolytics are excluded from the effective sample; reasons that are system related (e.g., equipment-related, staff-related, etc.) remain in the effective sample.

Results of the survey of the EWG also support the face validity of the exclusions and exceptions for NQF #0288, and indicate that these exclusions are consistent with prevailing gold standards of care or are necessary to support measure calculation.

2b4. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b4.1. What method of controlling for differences in case mix is used?
☒ **No risk adjustment or stratification**
☐ **Statistical risk model with** Click here to enter number of factors **risk factors**
☐ **Stratification by** Click here to enter number of categories **risk categories**
☐ **Other,** Click here to enter description

2b4.2. If an outcome or resource use measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

N/A

This measure is a process measure for which we provide no risk adjustment or stratification. We determined risk adjustment and stratification were not appropriate based on the measure evidence base and the measure construct. As a process-of-care measure, the decision to administer fibrinolytic therapy within 30 minutes of ED arrival when clinical guidelines are met should not be influenced by SDS factors; rather, adjustment would potentially mask such important inequities in care delivery. Variation across patient populations is reflective of differences in the quality of care provided to the disparate patient population included in the measure's denominator.

During the measure maintenance process, we perform an annual review of the literature, to identify articles and clinical practice guidelines published in the last 12 months, which includes a scan for potential patient subpopulations for which there are differences in the clinical decision to administer fibrinolytic therapy; this most recent review identified no clear evidence of an empirical relationship between SDS and timely administration of fibrinolytic therapy.

In addition to the evidence gathered from the literature, stakeholder feedback obtained during the eight years of implementation and public reporting has not identified concerns related to SDS factors and need for risk adjustment. This supports the conceptual model upon which the measure is based.

2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or sociodemographic factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*)

N/A

Not applicable - No risk adjustment or stratification was performed.

2b4.4a. What were the statistical results of the analyses used to select risk factors?

Not applicable - No risk adjustment or stratification was performed.

2b4.4b. Describe the analyses and interpretation resulting in the decision to select SDS factors (e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects)

Not applicable - No risk adjustment or stratification was performed.

2b4.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)
Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

Not applicable - No risk adjustment or stratification was performed.

2b4.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

Not applicable - No risk adjustment or stratification was performed.

2b4.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

Not applicable - No risk adjustment or stratification was performed.

2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

Not applicable - No risk adjustment or stratification was performed.

2b4.9. Results of Risk Stratification Analysis:

Not applicable - No risk adjustment or stratification was performed.

2b4.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)

Not applicable - No risk adjustment or stratification was performed.

2b4.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

Not applicable - No risk adjustment or stratification was performed.

_____

2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b5.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

N/A

Differences in performance scores for facilities meeting public-reporting requirements were tested. For the April 2014-March 2015 data collection period, this included 76 facilities. Additional details of this analysis are provided in **Section 2b5.2**.

2b5.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

After trending quarterly data for both national performance and benchmark performance, from Q4-08 to Q1-10, we have seen the following results: the measure has shown a constant gap in performance between the national rate and the benchmark rate since Q4-08.
Q1 2010 Analysis Provider Level
670 hospitals submitted 1,479 eligible cases.
Min 0
10th percentile 0
25th percentile 0
Median 50
75th percentile 100
90th percentile 100
Max 100
670 hospitals submitted 1,479 eligible cases. National rate: 53.5
Top 10% represented by benchmark results: 43 hospitals submitted 191 cases. Benchmark Rate: 98.4

Distribution of Facility Performance Scores

Mean	Std. Dev.	Min.	10th Percent	Lower Quartile	Median	Upper Quartile	90th Percent	Max.
70.6	19.5	9.0	42.0	58.0	73.0	84.5	93.0	100.0

2b5.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

The measure was able to detect facilities with better and worse than average performance. The facility performance scores ranged from 9.0% to 100.0%, with a median of 73.0%. Fifty percent of facilities fell within the interquartile range of 58.0%to 84.5%. The mean ± SD facility performance score was 70.6% ± 19.5%. Analysis of the April 2014-March 2015 performance data demonstrates the ability of the measure to identify outlying performance. By reporting a measure mean (benchmark value), this provides an opportunity for outlying facilities to identify underperformance related to timely administration of fibrinolytic therapy in situations when it is clinically appropriate.

2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS
If only one set of specifications, this section can be skipped.

Note: *This item is directed to measures that are risk-adjusted (with or without SDS factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without SDS factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b6.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

N/A

Not Applicable - this measure only uses one set of specifications.

2b6.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

N/A

Not Applicable - this measure only uses one set of specifications.

2b6.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

N/A

Not Applicable - this measure only uses one set of specifications.

2b7. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b7.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

This measure is calculated using chart-abstracted data. To limit the effects of missing data, abstractors cannot submit a value of "missing" for individual data elements. When they submit a value of "missing" the case is rejected from the abstraction tool. While abstractors cannot submit missing data, they may submit a value of "UTD" for select data elements for which missing information may be more likely, such as *Fibrinolytic Administration Time*. Cases where a value of "UTD" affects clinical decision making are excluded from the measure. Cases where a value of "UTD" is reflective of poor documentation are included in the effective sample but excepted from the numerator. To identify the extent and distribution of cases with a value of "UTD" for a data element, we calculated the frequency of such cases as well as the distribution of cases across eligible facilities. The frequency and distribution of missing data are described in **Section 2b3.3.**

2b7.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

The frequency and distribution of missing data are described in **Section 2b3.3**. We did not perform statistical analyses of missing data.

2b7.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i*.e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

As described in **Section 2b3.3**, the removal of cases from the effective sample and/or numerator where an abstractor submits a value of "UTD" are necessary to align with clinical guidelines and enable measure calculation. Additionally, these exclusions/exceptions limit the biasing effects of missing data. Cases where a value of "UTD" affects clinical decision-making are excluded from the measure. Cases where a value of "UTD" is reflective of poor documentation are included in the denominator but excepted from the numerator. Overall, 22 cases of the 3,844 cases in the effective sample (0.0%) have "UTD" value for the three numerator exception criteria, suggesting that such cases have a negligible effect on measure scores. The frequency and distribution of numerator exceptions are discussed in **Section 2b3.2**. This exclusion/exception approach penalizes facilities for poor documentation, but does not artificially include cases where administration of fibrinolytic therapy may not reflect appropriate care.

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)
If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. **To what extent are the specified data elements available electronically in defined fields?** *(i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields)*
Some data elements are in defined fields in electronic sources

3b.2. **If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources.**
NQF #0288 shares key data elements with an electronic clinical quality measure (eCQM), #0164: Fibrinolytic Therapy Received Within 30 Minutes of Hospital Arrival. The potential for e-specification will require special attention to the Initial ECG Interpretation and Reason for Delay in Fibrinolytic Therapy data elements, since these currently rely on logic and inferences that abstractors have been trained to interpret. Abstractors often rely on ECG print-outs and medical notes to determine the presence of ST-segment elevation on an ECG, and acceptable reasons for delay in fibrinolytic therapy typically require a review of medical notes.

Use of EHR data will require vendors to develop mechanisms to capture ECG findings, which are currently used to define the measure's denominator, in a structured field; because of the complicated nature of ECG interpretation, results included in a structured field should come from a review by an eligible provider rather than the ECG machine´s output. Because clinical or patient-centered reasons for not administering fibrinolytic therapy exist and effectively exclude patients from the measure, this data will also need to be captured and made available for measurement or clinical decision support within the EHR workflow.

The AMI and Stroke expert work group (EWG) considers NQF #0288 to be wholly feasible as it is currently specified, but considers e-specification to be moderately feasible. They concur that the key data elements for NQF #0288 are not readily available in a structured format within all EHR systems. In particular, EHR systems may need new structured fields for Initial ECG Interpretation and Reason for Delay in Fibrinolytic Therapy, which are not perceived to be a standard feature for most systems at this time.

Based on EWG feedback, the availability of the information from the data elements is highly dependent upon the EHR system used in each facility. If they cannot be translated into structured fields, then the data elements must be manually chart abstracted.

3b.3. **If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL.**
No feasibility assessment **Attachment:**

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.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 and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.**
IF a PRO-PM, consider implications for both individuals providing PROM data (patients, service recipients, respondents) and those whose performance is being measured.
An online survey of five members of the AMI and Stroke EWG with expertise in cardiology, neuro-radiology, emergency medicine, emergency transport, and emergency nursing was conducted to assess the face validity, feasibility, use, and usability of NQF #0228. All participants agreed or strongly agreed that patients who receive fibrinolytic therapy within 30 minutes of arrival to the ED can be accurately captured using chart-abstracted data. Additionally, the majority of participants agreed that practical aspects of reporting this chart-abstracted measure do not place undue burden on facilities that collect the data.

3c.2. **Describe any fees, licensing, or other requirements to use any aspect of the measure as specified** *(e.g., value/code set, risk model, programming code, algorithm)*.
No fees, licensure, or other requirements are necessary to use this measure; however, CPT codes, descriptions, and other data are

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Planned	Current Use (for current use provide URL)
	Public Reporting Hospital Outpatient Quality Reporting (HOQR) http://www.medicare.gov/hospitalcompare/search.html Hospital Outpatient Quality Reporting https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1192804531207 Quality Improvement with Benchmarking (external benchmarking to multiple organizations) Hospital Outpatient Quality Reporting Hospital Outpatient Quality Reporting http://www.medicare.gov/hospitalcompare/search.html

4a.1. For each CURRENT use, checked above, provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included

Public Reporting:
Name of program and sponsor: CMS HOQR Program
Purpose: The HOQR Program is a pay for quality data reporting program implemented by CMS for outpatient hospital services. In addition to providing hospitals with a financial incentive to report their quality of care measure data, the HOQR Program provides CMS with data to help Medicare beneficiaries make more informed decisions about their health care. Hospital quality of care information gathered through the HOQR Program is publicly available on the Hospital Compare website.
Accountable entities and patients: The publicly reported values (on Hospital Compare) are calculated for all facilities in the United States that meet minimum case count requirements. The number of facilities that met minimum case count criteria during the April 2010-March 2015 data collection periods ranged from 76-121 facilities, annually. The number of facilities meeting minimum case count criteria by year is presented in Section 1b.2.Facilities eligible to report this measure are subject to the Outpatient Prospective Payment System (OPPS) guidelines.

Quality Improvement with Benchmarking (external benchmarking to multiple organizations):
Name of program and sponsor: CMS HOQR Program

Purpose: The HOQR Program is a pay for quality data reporting program implemented by CMS for outpatient hospital services. In addition to providing hospitals with a financial incentive to report their quality of care measure data, the data is publicly reported on the Hospital Compare Website. The data reported on Hospital Compare not only shows the hospital's score on the measure, but also provides state and national averages for the measure. This enables consumers to compare the hospital's performance to other facilities and determine if the facility is an outlier.
Accountable entities and patients: The publicly reported values (on Hospital Compare) are calculated for all facilities in the United States that meet minimum case count requirements. The number of facilities that met minimum case count criteria during the April 2010-March 2015 data collection periods ranged from 76-121 facilities, annually. The number of facilities meeting minimum case count criteria by year is presented in Section 1b.2. Facilities eligible to report this measure are subject to the OPPS guidelines.

4a.2. **If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons?** (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)
This measure is publicly reported.

4a.3. **If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement.** (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)
This measure is publicly reported.

4b. Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b.1. **Progress on Improvement. (Not required for initial endorsement unless available.)**
Performance results on this measure (current and over time) should be provided in 1b.2 and 1b.4. Discuss:

- **Progress (trends in performance results, number and percentage of people receiving high-quality healthcare)**
- **Geographic area and number and percentage of accountable entities and patients included**

Summary statistics of performance scores from the April 2010-March 2015 data collection periods are provided in Section 1b.2.

The median rate of fibrinolytic therapy administered within 30 minutes of ED arrival, given the patient had a principal diagnosis associated with AMI, had ST-segment elevation of the ECG closest to ED arrival, and received fibrinolytic therapy, has increased between the 2010 and 2015 data collection periods (65.0% to 73.0%). One hundred twenty-one facilities met minimum case count requirements during the April 2010-March 2011 data collection period and 76 facilities met minimum case count requirements during the April 2014-March 2015 data collection period. During the April 2010-March 2011 data collection period, there were 1,930 sampled cases where a patient had a principal diagnosis associated with AMI, had ST-segment elevation of the ECG closest to ED arrival, and received fibrinolytic therapy. Of those cases, 1,257 were administered fibrinolytic therapy within 30 minutes of ED arrival (65.1%). During the April 2014-March 2015 data collection period, there were 1,221 sample cases where a patient had a principal diagnosis associated with AMI, had ST-segment elevation of the ECG closest to ED arrival, and received fibrinolytic therapy. Of those cases, 871 were administered fibrinolytic therapy within 30 minutes of ED arrival (71.3%).
These cases reflect only a subset of the patients eligible for the measure. Dependent upon the facility's total case count, the facility may report all cases or a sample of cases; thus, the number of patients receiving high-quality healthcare as performance on the measure improves is larger than the number of cases captured by the measure.

4b.2. **If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.**
Not applicable, as there is demonstrated improvement in measure performance over time.

4c. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4c.1. **Were any unintended negative consequences to individuals or populations identified during testing; OR has evidence of**

unintended negative consequences to individuals or populations been reported since implementation? If so, identify the negative unintended consequences and describe how benefits outweigh them or actions taken to mitigate them.
Measure testing did not identify any unintended consequences. Similarly, no evidence of unintended consequences to individuals or populations has been reported by external stakeholders since its implementation. The potential for unintended consequences will continue to be monitored through an annual review of the literature as well as an ongoing review of stakeholder comments and inquiries.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures
Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.
Yes

5.1a. **List of related or competing measures (selected from NQF-endorsed measures)**
0163 : Primary PCI received within 90 minutes of hospital arrival
0290 : Median Time to Transfer to Another Facility for Acute Coronary Intervention

5.1b. **If related or competing measures are not NQF endorsed please indicate measure title and steward.**
Median Time to Fibrinolysis – Centers for Medicare and Medicaid Services (CMS)
Fibrinolytic Therapy Received Within 30 Minutes of Hospital Arrival – CMS

5a. Harmonization

The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. **If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):**
Are the measure specifications completely harmonized?
No

5a.2. **If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.**
NQF #0288 and NQF #0290 are both in the HOQR Program, and NQF #0163 is included in the Hospital Inpatient Quality Reporting (HIQR) Program as an electronically specified clinical quality measure (eCQM). The two measures use the same initial patient population – patients with AMI and ST-segment elevation on the ECG performed closest to emergency department arrival who are transferred from the emergency department to a short-term general hospital for inpatient care, or to a Federal healthcare facility. While the target populations are the same, the focus of the two measures is different. NQF #0288 focuses on the timely administration of fibrinolytic therapy and the focus of NQF #0290 is the timely transfer of patients who require PCI. Although NQF #0163 (used in the HIQR Program) is similar to NQF #0288 (HOQR), the two measures serve different target populations and purposes: NQF #0288 focuses on timely administration of fibrinolytic therapy, while NQF #0163 focuses on the timely initiation of PCI for a patient who arrives at a PCI-capable hospital. All three measures share a number of key data elements (i.e., Initial ECG Interpretation, Fibrinolytic Administration, and Arrival Time). The specifications for the three measures are generally aligned, where possible.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. **If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):**
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)
No competing measures that address both the same measure focus and target population as NQF #0288 were identified.

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

No appendix **Attachment:**

Contact Information

Co.1 **Measure Steward (Intellectual Property Owner):** Centers for Medicare & Medicaid Services
Co.2 **Point of Contact:** Megan, Hayden, megan.hayden@cms.hhs.gov, 410-786-1970-
Co.3 **Measure Developer if different from Measure Steward:** The Lewin Group
Co.4 **Point of Contact:** Colleen, McKiernan, Colleen.McKiernan@lewin.com, 703-269-5595-

Additional Information

Ad.1 **Workgroup/Expert Panel involved in measure development**
Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
The contractor has convened an EWG, which evaluates and provides feedback on measure-development and maintenance efforts for a set of five AMI and one stroke measure. Specifically, the EWG provides direction and feedback through all phases of project activities, including expansion of the measures to additional CMS quality reporting programs, updates to the current specifications of these six measures, review of quantitative testing results, feedback on qualitative testing questions (i.e., results of EWG member questionnaires), and support for endorsement of the measures by the National Quality Forum (NQF).

The following is a list of the contractor's EWG members:

Joseph P. Drozda, Jr., MD
TEP 2010; Mercy Health, Rep. of American College of Cardiology; Director of Outcomes Research

Mustapha Ezzeddine, MD
University of Minnesota Medical Center, Director, Stroke Program

T. Bruce Ferguson, Jr., MD, FACC
TEP 2010; Brody School of Medicine at ECU, Dept. of Cardiovascular Sciences, Professor of Surgery and Physiology

Joseph V. Messer, MD, MACC
TEP 2010; Rush University Medical Center, Rep. of American Medical Association, Professor of Medicine

Cathy Olson, MSN, RN
Emergency Nurses Association (ENA), Institute for Quality, Safety, and Injury Prevention, Director

David Seidenwurm, MD
American Society of Neuroradiology (ASNR); American College of Radiologists (ACR)

Stephen Traub, MD
TEP 2010; Mayo Clinic, Department of Emergency Medicine, Chair

Paul D. Varosy, MD, FACC, FAHA, FHRS
TEP 2010; VA Eastern Colorado Health Care System, Director of Cardiac Electrophysiology

Matt Zavadsky, MS-HPA
National Association of Emergency Medical Technicians (NAEMT)

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.2 **Year the measure was first released:** 2008
Ad.3 **Month and Year of most recent revision:** 01, 2016
Ad.4 **What is your frequency for review/update of this measure?** Annually
Ad.5 **When is the next scheduled review/update for this measure?** 01, 2017

Ad.8 **Additional Information/Comments:**

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 2906
De.2. **Measure Title:** Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)
Co.1.1. **Measure Steward:** PCPI Foundation
De.3. **Brief Description of Measure:** Percentage of patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12 month period who also have a prior MI or a current or prior LVEF <40% who were prescribed beta-blocker therapy
1b.1. **Developer Rationale:** For patients with coronary artery disease (CAD), beta-blockers are recommended for 3 years after myocardial infarction or acute coronary syndrome. Beta-blockers, particularly carvedilol, metoprolol succinate, or bisoprolol which have been shown to reduce risk of death, are recommended indefinitely for patients with CAD and LV systolic dysfunction. These agents have proven efficacy in reducing angina onset and improving the ischemic threshold during exercise. In patients who have suffered an MI, beta-blockers significantly reduce deaths and recurrent MIs. (1) Nonadherence to cardioprotective medications is prevalent among outpatients with CAD and can be associated with a broad range of adverse outcomes, including all-cause and cardiovascular mortality, cardiovascular hospitalizations, and the need for revascularization procedures.(2) This measure is intended to promote beta-blocker usage in select patients with CAD.

References:
1. Fihn SD, Gardin JM, Abrams J, Berra K, Blankenship JC, Dallas AP, Douglas PS, Foody JM, Gerber TC, Hinderliter AL, King SB III, Kligfield PD, Krumholz HM, Kwong RYK, Lim MJ, Linderbaum JA, Mack MJ, Munger MA, Prager RL, Sabik JF, Shaw LJ, Sikkema JD, Smith CR Jr, Smith SC Jr, Spertus JA, Williams SV. 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, 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:e44-164.
2. Gibbons RJ, Abrams J, Chatterjee K, Daley J, Deedwania PC, Douglas JS, Ferguson TB Jr., Fihn SD, Fraker TD Jr., Gardin JM, O'Rourke RA, Pasternak RC, Williams SV. ACC/AHA 2002 guideline update for the management of patients with chronic stable angina: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1999 Guidelines for the Management of Patients with Chronic Stable Angina). 2002. Available at: www.acc.org/clinical/guidelines/stable/stable.pdf

S.4. **Numerator Statement:** Patients who were prescribed beta-blocker therapy
S.7. **Denominator Statement:** All patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12 month period who also have a prior MI (within the past 3 years) or a current or prior LVEF <40%
S.10. **Denominator Exclusions:** Documentation of medical reason(s) for not prescribing beta-blocker therapy (eg, allergy, intolerance, other medical reasons)
Documentation of patient reason(s) for not prescribing beta-blocker therapy (eg, patient declined, other patient reasons)
Documentation of system reason(s) for not prescribing beta-blocker therapy (eg, other reasons attributable to the health care system)

De.1. **Measure Type:** Process
S.23. **Data Source:** Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record
S.26. **Level of Analysis:** Clinician : Group/Practice, Clinician : Individual

IF Endorsement Maintenance – Original Endorsement Date: Most Recent Endorsement Date:

New Measure -- Preliminary Analysis

Criteria 1: Importance to Measure and Report

1a. Evidence

1a. Evidence. The evidence requirements for a *process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured.

This measure is the new eMeasure version of measure 0070. The information provided for Evidence and Opportunity for Improvement is identical to that submitted for 0070. Measure 0070 was discussed during CV Phase 3 in 2015. The ratings for evidence and opportunity for improvement will automatically be assigned to this eMeasure without further discussion.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☐ **Yes** ☒ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Evidence Summary:

- The developer provided a diagram demonstrating that beta-blocker therapy leads to a reduced risk of death, reduced angina onset, improved ischemic threshold during exercise, and reduced recurrent MIs in patients with prior MIs.
- The developer provided one clinical guideline with two recommendations to support beta-blocker therapy in patients with coronary artery disease, prior MI, or current or previous heart failure. The 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease states:
  - Beta-blocker therapy should be started and continued for 3 years in all patients with normal LV function after MI or ACS. **Class I, Level of Evidence: B**
  - Beta-blocker therapy should be used in all patients with LV systolic dysfunction (EF ≤40%) with heart failure or prior MI, unless contraindicated. (Use should be limited to carvedilol, metoprolol succinate, or bisoprolol, which have been shown to reduce risk of death.) **Class I, Level of Evidence: A**
- The systematic review of the evidence provided was associated with the guideline, however, a summary of the quality, quantity, and consistency of the body of evidence provided was associated with the guideline.
- The developer provided an additional study conducted in 2011 as an update to the systematic review of the body of evidence.

Exception to evidence: N/A

Guidance from the Evidence Algorithm: N/A

2015 discussion: *In 2015, during the previous review of this eMeasure and the claims-based/registry version of the measure, the Committee agreed that the evidence provided demonstrates that beta-blocker therapy in patients with CAD leads to a reduced risk of death, reduced angina onset, improved ischemic threshold during exercise and reduced recurrent MIs in patients with prior MIs.*

2015 votes*: H-16; M-0; L-0; IE-0*

Questions for the Committee:

- *Is the Committee willing to accept the prior discussion on Evidence?*

Preliminary rating for evidence: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- The developer provided the following performance rates from the 2014 PQRS Experience Report:

Year	Average Performance Rate
2010	71.4%
2011	82.1%
2012	69.9%
2013	74.2%
2014	79.3%

- The developer provided additional rates of beta-blocker prescriptions among patients with CAD from 2008-2010 NCDR PINNACLE Registry® data analyzed by Maddox et al. (2013):

	Index Clinic Visit	w/in yr. following index visit
BB Rx Rate	73.3%	77.3%
Median Rx Rate	78.4%	79.4%
Range	35.2% - 100%	46.2% - 100%

- The developer also provided a rate of 86.4% for beta-blocker prescriptions among CAD patients following an MI based on an earlier analysis of 2008-2009 Pinnacle Registry data conducted by Chan et al. (2010).

Disparities:

- The developer did not provide disparities data from the claims-based, registry, or electronically specified measure.
- The developer noted that while this measure is included in several federal reporting programs, those programs have not yet made disparities data available to analyze and report.
- The developer provided data on disparities from Chan et al. (2010) using PINNACLE Registry data from 2009 that demonstrated uninsured patients were less likely to receive beta-blocker therapy after MI compared to those with private health insurance (73.3% vs. 80.5%).

2015 discussion: *In 2015, during the previous review of this eMeasure and the claims-based/registry version of the measure, the Committee agreed that there was an opportunity for improvement based on the data provided from the registry measure but expressed the importance of obtaining performance data to adequately evaluate this eMeasure against this criterion in the future.*

2015 votes*: H-4; M-12; L-0; I-0*

Questions for the Committee:

○ *Is the Committee willing to accept the prior discussion on Performance Gap?*

Preliminary rating for opportunity for improvement: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence to Support Measure Focus

Comments: **Clinical trials have demonstrated that BBs reduce mortality in the candidate population.

**The developer provided a diagram demonstrating that beta-blocker therapy leads to a reduced risk of death, reduced angina onset, improved ischemic threshold during exercise, and reduced recurrent MIs in patients with prior MIs.

The developer provided one clinical guideline with two recommendations to support beta-blocker therapy in patients with coronary artery disease, prior MI, or current or previous heart failure. The 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease states:

Beta-blocker therapy should be started and continued for 3 years in all patients with normal LV function after MI or ACS. Class I, Level of Evidence: B

Beta-blocker therapy should be used in all patients with LV systolic dysfunction (EF ≤40%) with heart failure or prior MI, unless contraindicated. (Use should be limited to carvedilol, metoprolol succinate, or bisoprolol, which have been shown to reduce risk of death.) Class I, Level of Evidence: A

**Evidence appears sufficient."

**Evidence is the 2012 Guideline for the Management of Stable Angina - recommendation for beta blocker therapy for post -MI patients is IB (at least one RCT) and LV dysfucntion IA (multiple RCTs) per the ACCF. The process has direct effect on outcome. Rate as Moderate evidence. The study added for this application was an observational study..

1b. Performance Gap

Comments: **There is a significant performance gap. Disparity data are not provided

**There is a performance GAP as noted by PQRS (79%) and PINNACLE Registry (79%, range 46-100) data.

**1. The most recent data presented are from 2014 and only from Pinnacle data. Given that this measure has been in effect for a while, it seems that the developers could have produced more recent data and also additional sources.

2. It is not clear that these are data from an EHR. The 2015 discussion specifically requested more data.

**Performance gap information was provided for 2010 -2014 from PQRS. There is opportunity for improvement but the numbers seem stagnant during the reporting period (69.9%-82.1% but up and down).

Yes I would accept the prior assessment of a performance gap (Mod)".

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

2a1. Reliability Specifications

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented.

Data source(s): electronic clinical data; electronic health record. This is an eMeasure.

Specifications:

- HQMF specifications for the eMeasure are included in the document set on SharePoint. See eMeasure Technical Advisor review below.
- The numerator includes patients who were prescribed beta-blocker therapy.
  - 'Prescribed' may include a prescription given to the patient for beta-blocker therapy at one or more visits in the measurement period OR patient already taking beta-blocker therapy as documented in current medication list.
  - For patients with prior MI: beta-blocker therapy includes any agent within the beta-blocker drug class. As of 2015, no recommendations or evidence are cited in current stable ischemic heart disease guidelines for preferential use of specific agents.

- For patients with prior LVEF <40%: beta-blocker therapy includes bisoprolol, carvedilol, or sustained release metoprolol succinate.
- The denominator includes all patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12 month period who also have a prior MI (within the past 3 years) or a current or prior LVEF <40%.
  - Prior Myocardial Infarction (MI) is limited to those occurring within the past 3 years.
  - Two or more encounters are required to establish that the eligible professional has an existing relationship with the patient.
- Denominator exclusions/exceptions included are:
  - Documentation of medical reason(s) for not prescribing beta-blocker therapy (eg, allergy, intolerance, other medical reasons)
  - Documentation of patient reason(s) for not prescribing beta-blocker therapy (eg, patient declined, other patient reasons)
  - Documentation of system reason(s) for not prescribing beta-blocker therapy (eg, other reasons attributable to the health care system)
- The developer included the following variables as recommended data elements for stratification: race, ethnicity, administrative sex, and payer.
- A calculation algorithm is included.
- Details for handling missing data are provided.

Questions for the Committee :

- *Are all the data elements clearly defined?*
- *Is the measure logic clear?*
- *Is it likely this measure is consistently implemented?*

eMeasure Technical Advisor(s) review:

Submitted measure is an HQMF compliant eMeasure	The submitted eMeasure specifications follow the industry accepted format for eMeasure (HL7 Health Quality Measures Format (HQMF)). **HQMF specifications** ☒ **Yes** ☐ **No**
Documentation of HQMF or QDM limitations	All components in the measure logic of the submitted eMeasure are represented using the HQMF and QDM;
Value Sets	The submitted eMeasure specifications uses existing value sets when possible and uses new value sets that have been vetted through the VSAC
Measure logic is unambiguous	Submission includes test results from a simulated data set demonstrating the measure logic can be interpreted precisely and unambiguously Demonstrated through Bonnie testing Submission also includes testing results from multiple EHRs, but includes no information to identify which EHRs, if the EHRs are ONC certified and when the testing took place.
Feasibility Testing	The submission contains a feasibility assessment that addresses data element feasibility and provides a sample of Bonnie testing to indicate that the measure logic is feasible. This is a legacy eMeasure included in the Meaningful Use program. The developer submitted Bonnie testing results to establish feasibility, and provided a scorecard and Bonnie testing process and results in the measure testing attachment.

2a2. Reliability Testing attachment

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers.

SUMMARY OF TESTING

Reliability testing level ☒ **Measure score** ☐ **Data element** ☐ **Both**

Reliability testing performed with the data source and level of analysis indicated for this measure ☒ **Yes** ☐ **No**

Method(s) of reliability testing:

- The dataset included a total of 2,762 physicians participating in PQRS and reporting on this measure via the EHR option from January 2014 through December 2014.
  - # physicians with all required data elements and minimum # of eligible patients (**10**): **473**
  - Average # of eligible patients: **54.1**
  - Range of eligible patients: **10 to 836**
  - Total # of patients: **25,605** *[These were the patients that were associated with physicians who had 10 or more patients eligible for this measure and remained after exceptions were removed.]*
- Reliability testing was conducted at the measure score level, using a beta-binomial model to assess the signal-to-noise ratio. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in performance. The higher the reliability score, the greater is the confidence with which one can distinguish the performance of one physician from another. This is an appropriate test for measure score reliability. A reliability of 0.70 is generally considered a minimum threshold for reliability.

Results of reliability testing :

- Reliability was **0.69** when evaluated at the minimum number of eligible patients (10) and **0.92** when evaluated at the average number of eligible patients (54.1).

Guidance from the Reliability Algorithm: Precise specifications (Box 1) →Empirical reliability testing (Box 2) → Computed performance scores for measure entities (Box 4) →Appropriate method used (Box 5) →High/moderate reliability statistic and scope (Box 6) →Moderate

Questions for the Committee:

- *Is the test sample adequate to generalize for widespread implementation?*
- *Do the results demonstrate sufficient reliability so that differences in provider performance can be identified?*

Guidance from the Reliability Algorithm: Precise specifications (Box 1)→Empirical reliability testing conducted (Box 2)→Reliability testing conducted with computed performance measure score (Box 4)→Signal-to-noise analysis of the measure score conducted (this is appropriate) (Box 5)→High/moderate certainty or confidence that the performance measure scores are reliable (Box 6)→Moderate

Preliminary rating for reliability: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

2b. Validity

2b1. Validity: Specifications

2b1. Validity Specifications. This section should determine if the measure specifications are consistent with the evidence.

Specifications consistent with evidence in 1a. ☒ **Yes** ☐ **Somewhat** ☐ **No**

Question for the Committee:

- *Are the specifications consistent with the evidence?*

2b2. Validity testing

2b2. Validity Testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

SUMMARY OF TESTING

Validity testing level ☐ **Measure score** ☐ **Data element testing against a gold standard** ☒ **Both**

Method of validity testing of the measure score:

- ☒ **Face validity only**
- ☐ **Empirical validity testing of the measure score**

Validity testing method:

Data element testing:

- The dataset included 2004 data from an academic general internal medicine clinic using a commercial EHR. A sample of 134 patient charts were selected using random sampling via automated EHR review. Patient-level data element validity testing was conducted by comparing the values obtained from electronic extraction from the EHR to those obtained by manual abstraction from the EHR by the trained investigator (Validity against the Gold Standard). This is an appropriate test for data element testing.
- The developer stated that charts selected for abstraction included patients aged 18 years and older with a visit, problem list, or medical history diagnosis of CAD. It is not clear if the denominator inclusion criteria was selected through automated review, manual review, or both.

Face Validity:

- Face validity of the measure score was systematically assessed using an expert panel of 12 members from the AMA-PCPI Measure Advisory Committee.

Validity testing results:

Data element testing results:

- The developers provided percent agreement for the 134 patients sampled via automated EHR review. Of the 134 patients, the automated EHR review detected 111 patients (**82.8%)** that met the numerator criteria.
  - An additional 10 patients were detected through comparison of automated and manual EHR review and the percent agreement increased to **90.3%.**
  - The developer noted that the discrepancies between the EHR automated review alone and the automated review plus manual review were due to two types of misclassification: failure to correctly identify performance of quality measures among true, eligible patients; and failure to correctly exclude patients.
- The developer provided percent agreement of one final overall computation for all patients. NQF guidance for eMeasures states that testing at the level of data elements requires that all critical data elements be tested. At a minimum the numerator, denominator, and exclusions/exceptions must be assessed and reported separately. In addition to percent agreement, statistical analyses such as sensitivity and specificity, positive predictive value, and negative predictive value are required.

Face Validity Results:

- **91.7%** (12) of the respondents either agreed or strongly agreed that this measure can accurately distinguish good and poor quality.

Questions for the Committee:

- *Are the test samples adequate to generalize for widespread implementation?*
- *Do the testing results demonstrate sufficient validity so that conclusions about quality can be made?*
- *Do you agree that the score from this measure as specified is an indicator of quality?*

2b3-2b7. Threats to Validity

2b3. Exclusions:

- The developer analyzed 2,717 exceptions from 2009 for frequency and variability across five physician offices using five different EHRs and found that:
  - 2,292 (84.4%) exceptions were medical reasons for not prescribing beta blocker therapy
  - 347 (12.8%) exceptions were patient reasons for not prescribing beta blocker therapy

- 78 (2.9%) were system reasons for not prescribing beta blocker therapy
- The developer did not provide the total number of physicians, average number of exceptions per physician, or the overall exception rate.

Questions for the Committee:

- *Are the results from the exceptions analysis a threat to validity?*
- *Are any patients or patient groups inappropriately excluded from the measure?*
- *Are the number of exclusions reasonable?*

2b4. Risk adjustment: **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

2b5. Meaningful difference *(can statistically significant and clinically/practically meaningful differences in performance measure scores can be identified)*:

- The developer calculated measures of tendency, variability, and dispersion based on the sample of 473 physicians using EHRs:
  - Mean performance rate: 0.57
  - Median performance rate: 0.53
  - Mode: 0.5
  - Standard deviation: 0.25
  - Min, Max: 0.00-1.00
  - IQR: 0.4 (0.8 – 0.4)

Question for the Committee:

- *Does this eMeasure identify statistically and clinically meaningful differences about quality among providers?*

2b6. Comparability of data sources/methods:

- Measure is not specified for more than one data source; comparability of data sources is not needed.

2b7. Missing Data

- The developer stated that data are not available to complete an analysis of missing data but did specify how missing data are handled.

Guidance from the Validity Algorithm: Specifications consistent with evidence (Box 1)→Potential threats to validity assessed (Box 2)→Empirical validity testing (Box 3)→Face validity testing (Box 4) and patient-level data element validity testing (Box 10)→EHR extraction of patient-level data elements compared against the gold standard and overall percent agreement provided. Sensitivity, specificity, PPV, NPV for minimum of numerator, denominator, exceptions required (Box 11)→Moderate (highest eligible rating is MODERATE)

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2d)

2a1. & 2b1. Specifications

Comments: **Specs are clear

**Reliability: Moderate

The data elements are clearly defined, the measure logic is clear and the measure should be consistently implemented.

Specifications:

HQMF specifications for the eMeasure are included in the document set on SharePoint. See eMeasure Technical Advisor review below.

The numerator includes patients who were prescribed beta-blocker therapy.

'Prescribed' may include a prescription given to the patient for beta-blocker therapy at one or more visits in the measurement period OR patient already taking beta-blocker therapy as documented in current medication list.

For patients with prior MI: beta-blocker therapy includes any agent within the beta-blocker drug class. As of 2015, no

recommendations or evidence are cited in current stable ischemic heart disease guidelines for preferential use of specific agents.
For patients with prior LVEF <40%: beta-blocker therapy includes bisoprolol, carvedilol, or sustained release metoprolol succinate.
The denominator includes all patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12 month period who also have a prior MI (within the past 3 years) or a current or prior LVEF <40%.
Prior Myocardial Infarction (MI) is limited to those occurring within the past 3 years.
Two or more encounters are required to establish that the eligible professional has an existing relationship with the patient.
Denominator exclusions/exceptions included are:
Documentation of medical reason(s) for not prescribing beta-blocker therapy (eg, allergy, intolerance, other medical reasons)
Documentation of patient reason(s) for not prescribing beta-blocker therapy (eg, patient declined, other patient reasons)
Documentation of system reason(s) for not prescribing beta-blocker therapy (eg, other reasons attributable to the health care system)
Reliability testing showed score of 0.7. No issues with specifications, definitions or coding.
**Numerator: For patients with prior MI, it appears that no credit will be given if the health record indicates ""beta-blocker"" but does not specify which one
Denominator: No assurances are given that the past 3 years of data will be available for all patients. There is potential for the denominator to be biased toward more recent events or providers who serve a stable population
Exclusions for patient reasons and health system reasons are left vague"
**The reliability testing was done (Beta-binomial) at was reported as acceptable (Moderate) 0.69 for the minimum # of patients (10) and 0.92 for the average # (around 57)
The data elements are fairly clearly defined but the feasibility information alludes to an interface template for each EHR having to be developed.
The logic is clear.
Once mapping and a template are created, it should be consistent in the electronic format. I am concerned about charts with missing variables being excluded from analysis. On the one hand, you can't calculate the performance with missing data but if the missing data is the LVEF in particular, this is a quality issue that needs to be detected and addressed.
2a2. Reliability Testing
Comments: **While suffering when the number of cases was low, reliability was adequate
**Reliability testing showed score of 0.7. No issues with specifications, definitions or coding.
**Reliability was tested with an adequate scope of patients, but only minimal results were presented. Reliability with the beta-binomial is dependent not only on physician sample size, but also on success rate. Developers only reported two sample sizes without referencing success rates.
A more informative set of results could easily have been provided, including a plot of the estimated underlying beta distribution overlaid with histograms for success rates, along with a histogram of physician reliability estimates. No results on whether top or bottom performers could be discriminated were presented.
**See 2a1 and 2b1 response
2b2. Validity Testing
Comments: **Validity was adequate
**Validity testing performed by a panel who reviewed 134 charts, numerator criteria correct 82% manual chart review detected additional 10 patients. Denominator not tested (or I could not find it).
There appear to be many exclusions which is a threat to validity but the measure is an indicator of quality.
**Testing was done on 134 patients from a 2004 dataset. The following statement was provided for results of validity testing: Of the 134 patients sampled via automated EHR review, 111 patients (82.8%) that met the numerator criteria were detected. Performance on the measure was calculated to be 90.3% through comparison of automated and manual EHR review. These statements are confusing – it is not clear whether they are addressing the validity of the performance; 111 patients out of 134 who met the numerator criteria is very high. No results are presented for data element testing. In addition, the sample was from one site in 2004; generalizability is questionable, especially given the recent explosion of EHRs and problems with implementation.
**There was a panel of 12 experts for face validity and 11 agreed or strongly agreed.
Percent agreement of automated EHR and manual EHR data retrieval was performed only 134 charts. Overall agreement was 90.3%"
2b3. Exclusions Analysis
2b4. Risk Adjustment/Stratification for Outcome or Resource Use Measures

2b5. Identification of Statistically Significant & Meaningful Differences In Performance
2b6. Comparability of Performance Scores When More Than One Set of Specifications
2b7. Missing Data Analysis and Minimizing Bias
Comments: **"There are no significant threats to validity.
There is no risk adjustment.
I wasn't clear about the exception rate
**Performance rates for individual measures are found to vary across the measure reporting and testing programs.
This is expected, in that the performance rates are derived from different data sources, different practice sites, and variation in both the program implementation of the measures and approaches to implementation of the measures at individual practices or by the physicians in the practice sites included in the testing project.
Variation in performance rates across practice sites suggests that the measure is able to differentiate among practices. In addition, no single relatively high value of performance for a measure should be used as an indication of that measure being "topped out", and hence no longer important or meaningful to measure
**2b3-7 Threats to validity: the exceptions are reasonable by type but the developer did not include the total number of MDs, av # of exceptios per MD or a mean exception rate. These would have been more meaningful.
2b4. No risk adjustment - okay
2b5. Meaningful difference - most providers seem to be grouped around 50%
2b7. Missing data - application states that if there are missing data elements the case is excluded from analysis. I think this skews the data toward the positive. For example, if critical information like LVEF is omitted, the case is not analyzed.".
Risk adjustment: not applicable.
Comparability of performance measures: NA.
Missing data/no response: no.

Criterion 3. Feasibility

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

The Bonnie testing tool and environment (see feasibility assessment attachment):

- 72 synthetic patient records were used to test the measure logic and value sets. Bonnie testing includes negative and positive testing of each data element in the measure. Positive testing ensures patients expected to be included in the measure are included. Negative testing ensures that patients who do not meet the data criteria are not included in the measure. The following data elements and logic statement were tested to confirm actual results met expectations:
    - Age
    - Encounter
    - Diagnosis: Coronary artery disease, No MI
    - Procedure: Cardiac Surgery
    - Population 1: Diagnostic Study, Result: Ejection fraction result <40%
    - Population 1: Diagnosis: LVSD
    - Population 1: Diagnosis: LVSD severity moderate or severe
    - Population 2: Diagnosis Resolved: MI
    - Medication, Order: Beta Blocker Therapy
    - Medication, Active: Beta Blocker Therapy
    - Exceptions
- Bonnie tool testing results reached 100% coverage and confirmed there was a test case for each pathway of logic (negative and positive test cases).
- The measure also had a 100% passing rate which confirmed that all the test cases performed as expected.
- The developer stated that five physician offices using five different EHRs were included in the feasibility assessment; the name of the EHRs was not provided.

Questions for the Committee:

- *Are the required data elements routinely generated and used during care delivery?*
- *Are the required data elements available in electronic form, e.g., EHR or other electronic sources?*
- *Is the data collection strategy ready to be put into operational use?*
- *Does the eMeasure Feasibility Score Card demonstrate acceptable feasibility in multiple EHR systems and sites?*

Preliminary rating for feasibility: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 3: Feasibility

3a. Byproduct of Care Processes

3b. Electronic Sources

3c. Data Collection Strategy

Comments: **Feasibility has been demonstrated in 5 different EHRS.
**Feasibility: Moderate to Low (concern over assessment below that variation supports performance improvement need vs issues with data collection via EHR). See below.
Comments: Five physician offices, using five different Electronic Health Records (EHRs) were included in this feasibility assessment. All data elements were successfully able to be identified and exported to a warehouse for measure calculation by all 5 sites, included in the project. Each of the five practice sites mapped the data elements required for the CAD measure to their individual EHR. Since each practice had a unique set of data fields, at the time of this feasibility assessment, individual mapping of the data elements at the practice level was required. An interface template was also developed for each practice EHR which contained the unique set of data fields, validation requirements and acceptable values associated with the measure. To assure consistent capture of data across a disperse set of EHR systems, the interface template identified the submission of the prescribed coding system and standardized medical vocabulary as defined per the measure.
Since the conclusion of this measure testing project, value sets have been developed to capture data for this measure in a more standardized manner. No modifications have been made to the measure, as the measure is currently feasible for implementation, as specified.
Bonnie information on feasibility
Bonnie test cases were developed, in 2015, based on the latest HQMF R2 release. The measure logic performs as expected in the Bonnie system.
Feasibility score cards and simulated test cases are attached, for review.
**Only potential concerns relate to existence of data elements for historical data with sufficient date detail. Other data elements should be available.
**Used Bonnie platform to test - 100%
Not sure all charts include readily available time since infarction information in a readily retrievable manner.
Template mapping for specific EHRs seems to be necessary to have the data in fields where it can be easily retrieved. This is not a registry indicator. It is intended for private offices. They used PQRS as their test group -- an incentivized sample.

Criterion 4: Usability and Use

4. Usability and Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

Current uses of the measure

Publicly reported? ☐ **Yes** ☒ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No**

Accountability program details:

- Meaningful Use Stage 2 (EHR Incentive Program) (CMS): The Medicare and Medicaid EHR Incentive Programs provide incentive payments to eligible professionals, eligible hospitals, and critical access hospitals (CAHs) as they adopt, implement, upgrade or demonstrate meaningful use of certified EHR technology.
- Physician Quality Reporting System (PQRS) (CMS): The registry/claims-based version of this measure is in PQRS.

PQRS is a national reporting program that uses a combination of incentive payments and payment adjustments to promote reporting of quality information by eligible professionals (EPs). The program provides an incentive payment to practices with Eps.

Improvement results:

- The developer provided average performance rates from the registry/claims-based measure reported in PQRS.

Potential harms:

- None reported.

***Questions for the Committee**:*

- *How can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Do the benefits of the eMeasure outweigh any potential unintended consequences?*

Preliminary rating for usability and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 4: Usability and Use

4a. Accountability and Transparency

4b. Improvement

4c. Unintended Consequences

Comments: **It is in use in PQRS. It is not being publicly reported

**Usability and Use: Moderate

This is not being used in Public Reporting but it is being used ion accountability programs such as EHR Incentive Program and PQRS.

No unintended consequences from this measure.

**Given ambiguity regarding other criteria, it is not clear that this measure as described and implemented is reliable and valid.

**Not currently publically reported.

Is part of PQRS

Requires more than just claims data. Needs specific adaptations for each EHR and there are a lot of them out there."

Criterion 5: Related and Competing Measures

Related or competing measures:

- 0070 : Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)
- 0071 : Persistence of Beta-Blocker Treatment After a Heart Attack
- 0083 : Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)
- 2908 : Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)

Harmonization:

- Measure 0070 is the registry version of this eMeasure and is completely harmonized.
- Measure 2906 is harmonized with measures 0071, 0083, and 2908 to the extent possible. As a result, the denominator specifications for the measures differ where needed based on the differing patient populations.

Pre-meeting public and member comments

-

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0070

Measure Title: Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: 6/30/2015

Instructions

- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- Respond to all questions as instructed with answers immediately following the question. All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 10 pages (*incudes questions/instructions*; minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- Health outcome: [3] a rationale supports the relationship of the health outcome to processes or structures of care. Applies to patient-reported outcomes (PRO), including health-related quality of life/functional status, symptom/symptom burden, experience with care, health-related behavior.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- Efficiency: [6] evidence not required for the resource use component.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the U.S. Preventive Services Task Force (USPSTF) grading definitions and

methods, or Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.

5. 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 multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)
Outcome

- ☐ Health outcome: Click here to name the health outcome
- ☐Patient-reported outcome (PRO): Click here to name the PRO
  PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors
- ☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☒ Process: Beta-Blocker Therapy for CAD patients with Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)

☐ Structure: Click here to name the structure

☐ Other: Click here to name what is being measured

HEALTH OUTCOME/PRO PERFORMANCE MEASURE *If not a health outcome or PRO, skip to 1a.3*

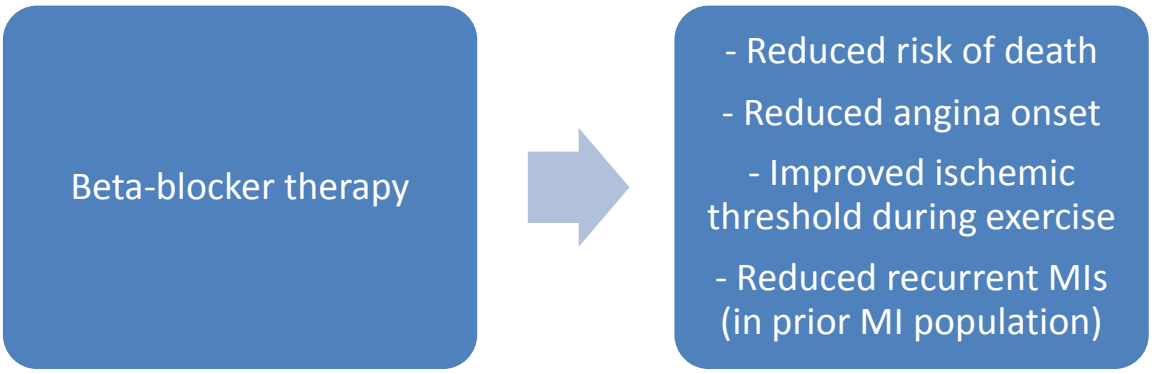
1a.2. Briefly state or diagram the path between the health outcome (or PRO) and the healthcare structures, processes, interventions, or services that influence it.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process, intervention, or service (*i.e., influence on outcome/PRO*).

Note: For health outcome/PRO performance measures, no further information is required; however, you may provide evidence for any of the structures, processes, interventions, or service identified above.

INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

1a.3. Briefly state or diagram the path between structure, process, intermediate outcome, and health outcomes. Include all the steps between the measure focus and the health outcome.



1a.3.1. What is the source of the systematic review of the body of evidence that supports the performance measure?

☒ Clinical Practice Guideline recommendation – ***complete sections 1a.4, and 1a.7***

☐ US Preventive Services Task Force Recommendation – ***complete sections 1a.5 and 1a.7***

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*) – ***complete sections 1a.6 and 1a.7***

☐ Other – ***complete section 1a.8***

Please complete the sections indicated above for the source of evidence. You may skip the sections that do not apply.

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION

1a.4.1. Guideline citation (*including date*) and **URL for guideline** (*if available online*):

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, 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:e44-164.

Available at: http://content.onlinejacc.org/article.aspx?articleid=1391404

1a.4.2. Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

4.4.2.2. BETA-BLOCKER THERAPY:

Beta-blocker therapy should be started and continued for 3 years in all patients with normal LV function after MI or ACS. (Class I, Level of Evidence: B)

Beta-blocker therapy should be used in all patients with LV systolic dysfunction (EF ≤40%) with heart failure or prior MI, unless contraindicated. (Use should be limited to carvedilol, metoprolol succinate,or bisoprolol, which have been shown to reduce risk of death.) (Class I, Level of Evidence: A)

1a.4.3. Grade assigned to the quoted recommendation with definition of the grade:

Both recommendation statements included in section 1a.4.2 have been assigned a Class I recommendation. Class I recommendations refer to " Conditions for which there is evidence and/or general agreement that a given procedure or treatment is beneficial, useful, and effective."

1a.4.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: If separate grades for the strength of the evidence, report them in section 1a.7.*)

ACCF/AHA guideline methodology categorizes indications as class I, II, or III on the basis of a multifactorial assessment of risk and expected efficacy viewed in the context of current knowledge and the relative strength of this knowledge. These classes summarize the recommendations for procedures or treatments as follows and noted in the table below:

Class I: Conditions for which there is evidence and/or general agreement that a given procedure or treatment is useful and effective.

Class II: Conditions for which there is conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of a procedure or treatment.

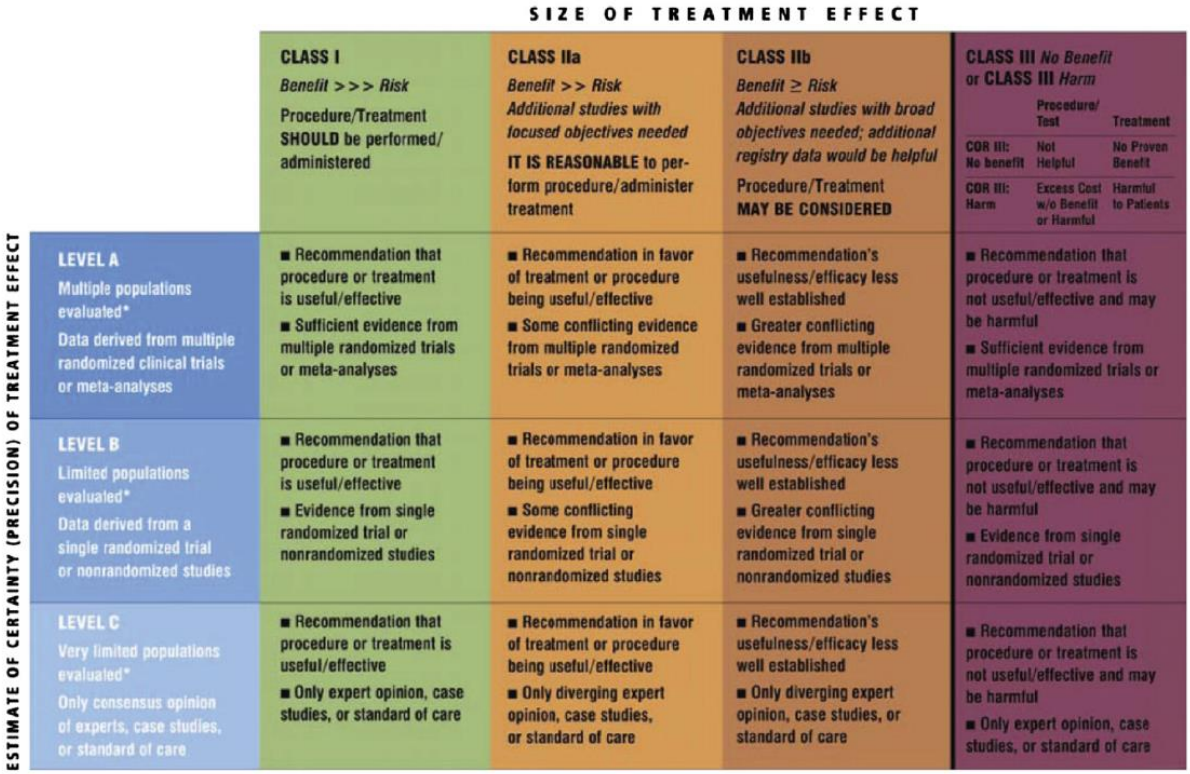
- IIa: Weight of evidence/opinion is in favor of usefulness/efficacy
- IIb: Usefulness/efficacy is less well established by evidence/opinion.

Class III: Conditions for which there is evidence and/or general agreement that the procedure/treatment is not useful/effective e and in some cases may be harmful.

- No Benefit- Procedure/Test not helpful or Treatment w/o established proven benefit
- Harm- Procedure/Test leads to excess cost w/o benefit or is harmful, and or Treatment is harmful

Additional detail regarding the classification of recommendation and level of evidence is provided in the following table:

Table 1. Applying Classification of Recommendation and Level of Evidence



ESTIMATE OF CERTAINTY (PRECISION) OF TREATMENT EFFECT	SIZE OF TREATMENT EFFECT: CLASS I	CLASS IIa	CLASS IIb	CLASS III *No Benefit* or CLASS III *Harm*
	Benefit >>> Risk Procedure/Treatment **SHOULD** be performed/ administered	*Benefit >> Risk* *Additional studies with focused objectives needed* **IT IS REASONABLE** to perform procedure/administer treatment	*Benefit ≥ Risk* *Additional studies with broad objectives needed; additional registry data would be helpful* Procedure/Treatment **MAY BE CONSIDERED**	Procedure/Test / Treatment COR III: No benefit — Not Helpful / No Proven Benefit COR III: Harm — Excess Cost w/o Benefit or Harmful / Harmful to Patients
LEVEL A Multiple populations evaluated* Data derived from multiple randomized clinical trials or meta-analyses	■ Recommendation that procedure or treatment is useful/effective ■ Sufficient evidence from multiple randomized trials or meta-analyses	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Sufficient evidence from multiple randomized trials or meta-analyses
LEVEL B Limited populations evaluated* Data derived from a single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is useful/effective ■ Evidence from single randomized trial or nonrandomized studies	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Evidence from single randomized trial or nonrandomized studies
LEVEL C Very limited populations evaluated* Only consensus opinion of experts, case studies, or standard of care	■ Recommendation that procedure or treatment is useful/effective ■ Only expert opinion, case studies, or standard of care	■ Recommendation in favor of treatment or procedure being useful/effective ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation's usefulness/efficacy less well established ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Only expert opinion, case studies, or standard of care

1a.4.5. Citation and URL for methodology for grading recommendations (*if different from 1a.4.1*)**:**

ACCF/AHA Task Force on Practice Guidelines. Methodology Manual and Policies From the ACCF/AHA Task Force on Practice Guidelines. American College of Cardiology Foundation and American Heart Association, Inc. Cardiosource.com. 2010. Available at: http://assets.cardiosource.com/Methodology_Manual_for_ACC_AHA_Writing_Committees.pdf and http://my.americanheart.org/idc/groups/ahamah-public/@wcm/@sop/documents/downloadable/ucm_319826.pdf

1a.4.6. If guideline is evidence-based (rather than expert opinion), are the details of the quantity, quality, and consistency of the body of evidence available (e.g., evidence tables)?

☐ Yes → ***complete section 1a.7***

☒ No → ***report on another systematic review of the evidence in sections 1a.6 and 1a.7; if another review does not exist, provide what is known from the guideline review of evidence in 1a.7***

1a.5. UNITED STATES PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

1a.5.1. Recommendation citation (*including date*) and **URL for recommendation** (*if available online*):

1a.5.2. Identify recommendation number and/or page number and **quote verbatim, the specific recommendation**.

1a.5.3. Grade assigned to the quoted recommendation with definition of the grade:

1a.5.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: the grading system for the evidence should be reported in section 1a.7.*)

1a.5.5. Citation and URL for methodology for grading recommendations (*if different from 1a.5.1*)**:**

Complete section 1a.7

1a.6. OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1. Citation (*including date*) and **URL** (*if available online*):

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

Complete section 1a.7

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

If more than one systematic review of the evidence is identified above, you may choose to summarize the one (or more) for which the best information is available to provide a summary of the quantity, quality, and consistency of the body of evidence. Be sure to identify which review is the basis of the responses in this section and if more than one, provide a separate response for each review.

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

The section of the guideline which includes the recommendations referenced in 1a4.2. pertains to the use of beta-blocker therapy in patients with stable ischemic heart disease

1a.7.2. Grade assigned for the quality of the quoted evidence with definition of the grade:

The guideline recommendations refer to 2 distinct patient populations address by the measure – 1) patients with a prior (resolved) (within the past 3 years) myocardial infarction and 2) patients with left ventricular systolic dysfunction (LVEF <40%). For the prior MI population, the weight of the evidence in support of the recommendation is rated as Level B. Level B evidence refers to "Data derived from a single randomized trial, or nonrandomized studies." For the LVSD population, the weight of the evidence in support of the recommendation is rated as Level A. Level A evidence refers to "Data derived from multiple randomized clinical trials or meta-analyses."

1a.7.3. Provide all other grades and associated definitions for strength of the evidence in the grading system.

Levels A and B are described in 1a.7.2. Level C evidence refers to "Only consensus opinion of experts, case studies, or standard-of-care." Additional details and information about the evidence rating scheme can also be seen in 1a.4.2. and 1a.4.3.

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range: 1996-2009

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

Information regarding the total number of studies and type of study designs included in the body of evidence is not available.

However, for the prior MI population: the guideline cites 3 articles in support of the recommendation statement. They include 2 systematic reviews including 33 and 82 randomized controlled trials, respectively, dating back to 1980. The third article was an observational study.

For the LVSD population: the guideline cites 5 articles in support of the recommendation statement. They include 3 randomized controlled trials, 1 meta-analysis of randomized controlled trials and 1 comparative analysis of randomized controlled trials.

1a.7.6. What is the overall quality of evidence <u>across studies</u> in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

Information regarding the overall quality of evidence across studies is not available.

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) <u>across studies</u> in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

The guideline does not include an overall estimate of benefit from the body of evidence. However, they do include the following summary information regarding the benefits of beta-blocker therapy, "Decreases in the rate–BP product, AV nodal conduction, and myocardial contractility from beta blockers reduce myocardial oxygen demand, counteracting beta-receptor activity and contributing to a reduction in angina onset, with improvement in the ischemic threshold during exercise and in symptoms. These agents significantly reduce deaths and recurrent MIs in patients who have suffered a MI and are especially effective when a STEMI is complicated by persistent or recurrent ischemia or tachyarrhythmias early after the onset of infarction."

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

The guideline describes the principle adverse effects of beta blockers as fatigue, exercise intolerance, lethargy, insomnia, nightmares, and impotence.

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for <u>each</u> new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

The guidelines reviewed and incorporated relevant new clinical trials published in peer-reviewed journals and articles through December 2011. A Medical Subject Headings (MeSH®) was conducted using the terms "Adrenergic beta-Antagonists"[Mesh] AND "Coronary Artery Disease"[Mesh] to identify articles published after 2011, resulting in 75 articles.

The articles that are most relevant to the focus of the body of evidence are described below.

1. Citation: Bangalore S1, Steg G, Deedwania P, Crowley K, Eagle KA, Goto S, Ohman EM, Cannon CP, Smith SC, Zeymer U, Hoffman EB, Messerli FH, Bhatt DL; REACH Registry Investigators. β-Blocker use and clinical outcomes in stable outpatients with and without coronary artery disease. JAMA. 2012 Oct 3;308(13):1340-9. doi: 10.1001/jama.2012.12559.

Description: Longitudinal, observational study of patients in the Reduction of Atherothrombosis for Continued Health (REACH) registry who were divided into 3 cohorts: known prior MI (n = 14,043), known CAD without MI (n = 12,012), or those with CAD risk factors only (n = 18,653) to assess the association of β-blocker use with cardiovascular events in stable patients with a prior history of MI, in those with CAD but no history of MI, and in those with only risk factors for CAD.

Results: With a median follow-up of 44 months (interquartile range, 35-45 months), event rates were not significantly different in patients with β-blocker use compared with those without β-blocker use for any of the outcomes tested, even in the prior MI cohort (489 [16.93%] vs 532 [18.60%], respectively; hazard ratio [HR], 0.90 [95% CI, 0.79-1.03]; P = .14). In the CAD without MI cohort, the associated event rates were not significantly different in those with β-blocker use for the

primary outcome (391 [12.94%]) vs without β-blocker use (405 [13.55%]) (HR, 0.92 [95% CI, 0.79-1.08]; P = .31), with higher rates for the secondary outcome (1101 [30.59%] vs 1002 [27.84%]; odds ratio [OR], 1.14 [95% CI, 1.03-1.27]; P = .01) and for the tertiary outcome of hospitalization (870 [24.17%] vs 773 [21.48%]; OR, 1.17 [95% CI, 1.04-1.30]; P = .01). In the cohort with CAD risk factors only, the event rates were higher for the primary outcome with β-blocker use (467 [14.22%]) vs without β-blocker use (403 [12.11%]) (HR, 1.18 [95% CI, 1.02-1.36]; P = .02), for the secondary outcome (870 [22.01%] vs 797 [20.17%]; OR, 1.12 [95% CI, 1.00-1.24]; P = .04) but not for the tertiary outcomes of MI (89 [2.82%] vs 68 [2.00%]; HR, 1.36 [95% CI, 0.97-1.90]; P = .08) and stroke (210 [6.55%] vs 168 [5.12%]; HR, 1.22 [95% CI, 0.99-1.52]; P = .06). However, in those with recent MI (≤1 year), β-blocker use was associated with a lower incidence of the secondary outcome (OR, 0.77 [95% CI, 0.64-0.92]).

Conclusion: Although this observational study found that the use of β-blockers in the populations studied was not associated with a lower risk of composite cardiovascular events, the article received several letters which highlighted 2 primary concerns: 1) the use of an observational study to assess the effectiveness of a drug when large RCTs and meta-analyses already have shown its effectiveness and 2) the study did not distinguish among different types of beta-blockers. As the measure developer, we would wait until an updated systematic review of the body of evidence is conducted which can confirm or refute the findings of the study taking into account the full body of evidence available.

_____

1a.8 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.8.1 What process was used to identify the evidence?

1a.8.2. Provide the citation and summary for each piece of evidence.

1. Evidence, Performance Gap, Priority – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all subcriteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form
evidence_attachment_CAD_BB-635917579089836457.docx

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- disparities in care across population groups.

1b.1. **Briefly explain the rationale for this measure** *(e.g., the benefits or improvements in quality envisioned by use of this measure)*
For patients with coronary artery disease (CAD), beta-blockers are recommended for 3 years after myocardial infarction or acute coronary syndrome. Beta-blockers, particularly carvedilol, metoprolol succinate, or bisoprolol which have been shown to reduce risk of death, are recommended indefinitely for patients with CAD and LV systolic dysfunction. These agents have proven efficacy in reducing angina onset and improving the ischemic threshold during exercise. In patients who have suffered an MI, beta-blockers significantly reduce deaths and recurrent MIs. (1) Nonadherence to cardioprotective medications is prevalent among outpatients with CAD and can be associated with a broad range of adverse outcomes, including all-cause and cardiovascular mortality, cardiovascular hospitalizations, and the need for revascularization procedures.(2) This measure is intended to promote beta-blocker usage in select patients with CAD.

References:

1. Fihn SD, Gardin JM, Abrams J, Berra K, Blankenship JC, Dallas AP, Douglas PS, Foody JM, Gerber TC, Hinderliter AL, King SB III, Kligfield PD, Krumholz HM, Kwong RYK, Lim MJ, Linderbaum JA, Mack MJ, Munger MA, Prager RL, Sabik JF, Shaw LJ, Sikkema JD, Smith CR Jr, Smith SC Jr, Spertus JA, Williams SV. 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, 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:e44-164.
2. Gibbons RJ, Abrams J, Chatterjee K, Daley J, Deedwania PC, Douglas JS, Ferguson TB Jr., Fihn SD, Fraker TD Jr., Gardin JM, O'Rourke RA, Pasternak RC, Williams SV. ACC/AHA 2002 guideline update for the management of patients with chronic stable angina: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1999 Guidelines for the Management of Patients with Chronic Stable Angina). 2002. Available at: www.acc.org/clinical/guidelines/stable/stable.pdf

1b.2. **Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.** *(This is required for endorsement maintenance. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included). This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*
2014 PQRS Experience Report
2014 is the most recent year for which PQRS Experience Report measure data is available. The average performance rates on Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%) over the last several years are as follows:
2010: 71.4%
2011: 82.1%
2012: 69.9%
2013: 74.2%
2014: 79.3%

It is important to note that PQRS has been and remains a voluntary reporting program. In the early years of the PQRS program,

participants received an incentive for satisfactorily reporting. However, beginning in 2015, the program imposes payment penalties for non-participants based on 2013 performance. 62% of eligible professionals participated using any reporting option in 2014. As a result, performance rates may not be nationally representative.

Reference: Center for Medicare and Medicaid Services. 2014 Reporting Experience Including Trends (2007-2014). Available at: https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/pqrs/analysisandpayment.html

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.
Suboptimal rates of beta-blocker prescriptions among patients with CAD indicated by PQRS data are further evidenced by several recent studies.

Maddox and colleagues analyzed data from 2008 through 2010 from the NCDR® PINNACLE Registry®, a national outpatient cardiology practice registry, to assess practice variation of secondary prevention medication prescription among CAD patients. Among eligible patients, beta-blockers were prescribed in 73.3% (63,800/86,999) at their index clinic visit. After inclusion of all visits among eligible patients occurring within the year following the index visit, the rates increased to 77.3%. Among practices, the median prescription rate of beta-blockers for eligible patients at their index clinic visit was 78.4% (range 35.2-100%) and 79.4% (range 46.2-100%) after inclusion of all visits among eligible patients occurring within the year following the index visit.(1)

An earlier study by Chan and colleagues analyzed 2008-9 data from the Pinnacle registry and found slightly higher rates (86.4%) of beta-blocker prescription among CAD patients following an MI. It's important to note that the Chan et al. study examined compliance rates with performance measures among the first 14,000 outpatients enrolled in the PINNACE program as compared to the Maddox et al study which included a larger and more heterogeneous patient and practice population.(2)

References:
1. Maddox TM, Chan PS, Spertus JA, Tang F, Jones P, Ho PM, Bradley SM, Tsai TT, Bhatt DL, Peterson PN. Variations in coronary artery disease secondary prevention prescriptions among outpatient cardiology practices: insights from the NCDR (National Cardiovascular Data Registry). J Am Coll Cardiol. 2014 Feb 18;63(6):539-46. doi: 10.1016/j.jacc.2013.09.053. Epub 2013 Oct 30.

2. Chan PS, Oetgen WJ, Buchanan D, et al. Cardiac performance measure compliance in outpatients: the American College of Cardiology and National Cardiovascular Data Registry´s PINNACLE (Practice Innovation And Clinical Excellence) program. J. Am. Coll. Cardiol. 2010; 56(1):8–14.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for endorsement maintenance. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*
While this measure is included in several federal reporting programs, those programs have not yet made disparities data available for us to analyze and report.

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations.
The Chan et al. article cited above conducted a secondary analysis of PINNACLE data for select performance measures to examine whether compliance rates differed by race or sex. The authors found that compliance rates were similar between black and white patients and men and women for all 4 CAD performance measures (including beta-blocker therapy after MI). (1)

A separate analysis was completed using PINNACLE data from 2009 to compare treatment rates by insurance status for 5 quality-of-care indicators for CAD care related to medication treatment. Uninsured patients were less likely to receive ß-blocker therapy after MI as compared with those who had private health insurance (73.3% vs. 80.5%; unadjusted RR=0.91; 95% CI, 0.87-0.95; P<0.001). There were no meaningful differences in treatment rates between patients with public and private insurance. (2)

1. Chan PS, Oetgen WJ, Buchanan D, et al. Cardiac performance measure compliance in outpatients: the American College of Cardiology and National Cardiovascular Data Registry´s PINNACLE (Practice Innovation And Clinical Excellence) program. J. Am. Coll. Cardiol. 2010; 56(1):8–14.

2. Smolderen KG, Spertus JA, Tang F, et al. Treatment Differences by Health Insurance Among Outpatients with Coronary Artery

Disease: Insights from the NCDR®. J Am Coll Cardiol. 2013 Mar 12; 61(10): 1069–1075.

1c. High Priority (previously referred to as High Impact)

The measure addresses:

- a specific national health goal/priority identified by DHHS or the National Priorities Partnership convened by NQF; OR
- a demonstrated high-priority (high-impact) aspect of healthcare (e.g., affects large numbers of patients and/or has a substantial impact for a smaller population; leading cause of morbidity/mortality; high resource use (current and/or future); severity of illness; and severity of patient/societal consequences of poor quality).

1c.1. **Demonstrated high priority aspect of healthcare**

Affects large numbers, A leading cause of morbidity/mortality, High resource use, Patient/societal consequences of poor quality, Severity of illness

1c.2. **If Other:**

1c.3. **Provide epidemiologic or resource use data that demonstrates the measure addresses a high priority aspect of healthcare. List citations in 1c.4.**

Heart disease is the leading cause of death for both men and women in the United States and coronary heart disease is the most common type of heart disease.(1) According to the American Heart Association's 2015 Heart Disease and Stroke Statistics, coronary heart disease alone caused ˜1 of every 7 deaths in the United States in 2011. In 2011, close to 400,000 Americans died of coronary heart disease. Each year, an estimated ˜635 000 Americans have a new coronary attack (defined as first hospitalized myocardial infarction or coronary heart disease death) and ˜300 000 have a recurrent attack. It is estimated that an additional 155 000 silent first myocardial infarctions occur each year. Approximately every 34 seconds, 1 American has a coronary event, and approximately every 1 minute 24 seconds, an American will die of one.(2)

1c.4. **Citations for data demonstrating high priority provided in 1a.3**

1. CDC, NCHS. Underlying Cause of Death 1999-2013 on CDC WONDER Online Database, released 2015. Data are from the Multiple Cause of Death Files, 1999-2013, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program.
2. Mozaffarian D, Benjamin EJ, Go AS, et al.; on behalf of the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics—2015 update: a report from the American Heart Association. Circulation. 2015;131:e29–e322.

1c.5. **If a PRO-PM (e.g. HRQoL/functional status, symptom/burden, experience with care, health-related behaviors), provide evidence that the target population values the measured PRO and finds it meaningful.** *(Describe how and from whom their input was obtained.)*

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the subcriteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. **Subject/Topic Area** *(check all the areas that apply):*

Cardiovascular, Cardiovascular : Ischemic Heart Disease, Coronary Artery Disease

De.6. **Cross Cutting Areas** *(check all the areas that apply):*

S.1. **Measure-specific Web Page** *(Provide a URL link to a web page specific for this measure that contains current detailed*

specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)
The measure specifications are included as an attachment with this submission. Additional measure details may be found at:http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/eCQM_Library.html Value sets at https://vsac.nlm.nih.gov

S.2a. **If this is an eMeasure**, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)
This is an eMeasure **Attachment:** PCPI_2906_CMS145v5_CAD-BB.zip

S.2b. **Data Dictionary, Code Table, or Value Sets** *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*
Attachment **Attachment:** PCPI_2906_CMS145_CAD-BB_ValueSets.xlsx

S.3. **For endorsement maintenance**, please briefly describe any changes to the measure specifications since last endorsement date and explain the reasons.
Supporting guidelines and specifications included in the measure are reviewed on an annual basis. The measure and specifications have been updated to align with the 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the Diagnosis and Management of Patients With Stable Ischemic Heart Disease which now recommends the use of beta-blocker therapy for 3 years following MI or ACS as opposed to indefinite use as previously recommended. Any additional changes have been incorporated during the annual update process to adhere to current eCQM industry standards while preserving the original measure intent.

S.4. **Numerator Statement** *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome)*
IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
Patients who were prescribed beta-blocker therapy

S.5. **Time Period for Data** *(What is the time period in which data will be aggregated for the measure, e.g., 12 mo, 3 years, look back to August for flu vaccination? Note if there are different time periods for the numerator and denominator.)*
Once during the 12 consecutive month measurement period

S.6. **Numerator Details** *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
For EHR:
HQMF eMeasure developed and is included in this submission.

We have provided the following definitions and/or guidance for convenience; please see HQMF eMeasure for complete details related to the specification.

NUMERATOR DEFINITION:
Prescribed may include prescription given to the patient for beta-blocker therapy at one or more visits in the measurement period OR patient already taking beta-blocker therapy as documented in current medication list.

NUMERATOR GUIDANCE:
Beta-blocker therapy:
- For patients with prior MI, beta-blocker therapy includes any agent within the beta-blocker drug class. As of 2015, no recommendations or evidence are cited in current stable ischemic heart disease guidelines for preferential use of specific agents
- For patients with prior LVEF <40%, beta-blocker therapy includes the following: bisoprolol, carvedilol, or sustained release metoprolol succinate

S.7. **Denominator Statement** *(Brief, narrative description of the target population being measured)*

All patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12 month period who also have a prior MI (within the past 3 years) or a current or prior LVEF <40%

S.8. **Target Population Category** *(Check all the populations for which the measure is specified and tested if any):*
Senior Care

S.9. **Denominator Details** *(All information required to identify and calculate the target population/denominator such as definitions, specific data collection items/responses , code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
For EHR:
HQMF eMeasure developed and is included in this submission.

We have provided the following definitions and/or guidance for convenience; please see HQMF eMeasure for complete details related to the specification.

DENOMINATOR DEFINITION:
Prior Myocardial Infarction (MI) for denominator 2 is limited to those occurring within the past 3 years.

DENOMINATOR GUIDANCE:
The requirement of "Count >=2 of Encounter, Performed" is to establish that the eligible professional has an existing relationship with the patient.

S.10. **Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*
Documentation of medical reason(s) for not prescribing beta-blocker therapy (eg, allergy, intolerance, other medical reasons)
Documentation of patient reason(s) for not prescribing beta-blocker therapy (eg, patient declined, other patient reasons)
Documentation of system reason(s) for not prescribing beta-blocker therapy (eg, other reasons attributable to the health care system)

S.11. **Denominator Exclusion Details** *(All information required to identify and calculate exclusions from the denominator such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
Exceptions are used to remove a patient from the denominator of a performance measure when the patient does not receive a therapy or service AND that therapy or service would not be appropriate due to patient-specific reasons. The patient would otherwise meet the denominator criteria. Exceptions are not absolute, and are based on clinical judgment, individual patient characteristics, or patient preferences. This measure was developed using the PCPI exception methodology which uses three categories of reasons for which a patient may be removed from the denominator of an individual measure. These measure exception categories are not uniformly relevant across all measures; for each measure, there must be a clear rationale to permit an exception for a medical, patient, or system reason. Examples are provided in the measure exception language of instances that may constitute an exception and are intended to serve as a guide to clinicians. For this measure, exceptions may include medical reason(s) (eg, allergy, intolerance, other medical reasons), patient reason(s) (eg, patient declined, other patient reasons) or system reason(s) for not prescribing beta-blocker therapy (eg, other reasons attributable to the health care system). Where examples of exceptions are included in the measure language, value sets for these examples are developed and included in the eSpecifications. Although this methodology does not require the external reporting of more detailed exception data, the PCPI recommends that physicians document the specific reasons for exception in patients' medical records for purposes of optimal patient management and audit-readiness. The PCPI also advocates the systematic review and analysis of each physician's exceptions data to identify practice patterns and opportunities for quality improvement.

Additional details by data source are as follows:
For EHR:
HQMF eMeasure developed and is included in this submission.

S.12. **Stratification Details/Variables** *(All information required to stratify the measure results including the stratification variables, definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b)*
Consistent with CMS' Measures Management System Blueprint and recent national recommendations put forth by the IOM and NQF to standardize the collection of race and ethnicity data, we encourage the results of this measure to be stratified by race, ethnicity, administrative sex, and payer and have included these variables as recommended data elements to be collected.

S.13. **Risk Adjustment Type** (Select type. Provide specifications for risk stratification in S.12 and for statistical model in S.14-15)
No risk adjustment or risk stratification
If other:

S.14. **Identify the statistical risk model method and variables** *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development and testing should be addressed with measure testing under Scientific Acceptability)*
No risk adjustment or risk stratification

S.15. **Detailed risk model specifications** *(must be in attached data dictionary/code list Excel or csv file. Also indicate if available at measure-specific URL identified in S.1.)*
Note: Risk model details (including coefficients, equations, codes with descriptors, definitions), should be provided on a separate worksheet in the suggested format in the Excel or csv file with data dictionary/code lists at S.2b.
Provided in response box S.15a

S.15a. **Detailed risk model specifications** (*if not provided in excel or csv file at S.2b*)
No risk adjustment or risk stratification

S.16. **Type of score:**
Rate/proportion
If other:

S.17. **Interpretation of Score** *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*
Better quality = Higher score

S.18. **Calculation Algorithm/Measure Logic** *(Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.)*
To calculate performance rates:
1. Find the patients who meet the initial population (ie, the general group of patients that a set of performance measures is designed to address).
2. From the patients within the initial population criteria, find the patients who qualify for the denominator (ie, the specific group of patients for inclusion in a specific performance measure based on defined criteria). Note: in some cases the initial population and denominator are identical.
3. From the patients within the denominator, find the patients who meet the numerator criteria (ie, the group of patients in the denominator for whom a process or outcome of care occurs). Validate that the number of patients in the numerator is less than or equal to the number of patients in the denominator
4. From the patients who did not meet the numerator criteria, determine if the provider has documented that the patient meets any criteria for exception when denominator exceptions have been specified [for this measure: medical reason(s) (eg, allergy, intolerance, other medical reasons), patient reason(s) (eg, patient declined, other patient reasons) or system reason(s) for not prescribing beta-blocker therapy (eg, other reasons attributable to the health care system).] If the patient meets any exception criteria, they should be removed from the denominator for performance calculation. --Although the exception cases are removed from the denominator population for the performance calculation, the exception rate (ie, percentage with valid exceptions) should be calculated and reported along with performance rates to track variations in care and highlight possible areas of focus for QI.

If the patient does not meet the numerator and a valid exception is not present, this case represents a quality failure.

S.19. **Calculation Algorithm/Measure Logic Diagram URL or Attachment** *(You also may provide a diagram of the Calculation Algorithm/Measure Logic described above at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
No diagram provided

S.20. **Sampling** *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*
IF a PRO-PM, identify whether (and how) proxy responses are allowed.
Not applicable. The measure is not based on a sample.

S.21. **Survey/Patient-reported data** *(If measure is based on a survey, provide instructions for conducting the survey and guidance on minimum response rate.)*
IF a PRO-PM, specify calculation of response rates to be reported with performance measure results.
Not applicable. The measure is not based on a survey.

S.22. **Missing data** (specify how missing data are handled, e.g., imputation, delete case.)
Required for Composites and PRO-PMs.
Patient eligibility is determined by a set of defined criteria relevant to a particular measure. If data required to determine patient eligibility are missing, those patients/cases would be ineligible for inclusion in the denominator and therefore the patient/case would be deleted.

If data required to determine if a denominator eligible patient qualifies for the numerator (or has a valid exclusion/exception) are missing, this case would represent a quality failure.

S.23. **Data Source** *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*
If other, please describe in S.24.
Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record

S.24. **Data Source or Collection Instrument** *(Identify the specific data source/data collection instrument e.g. name of database, clinical registry, collection instrument, etc.)*
IF a PRO-PM, identify the specific PROM(s); and standard methods, modes, and languages of administration.
Not applicable.

S.25. **Data Source or Collection Instrument** *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
No data collection instrument provided

S.26. **Level of Analysis** *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*
Clinician : Group/Practice, Clinician : Individual

S.27. **Care Setting** *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*
Ambulatory Care : Clinician Office/Clinic, Home Health, Other, Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility
If other: Domiciliary

S.28. **COMPOSITE Performance Measure** - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

2a. Reliability – See attached Measure Testing Submission Form
2b. Validity – See attached Measure Testing Submission Form
NQF_2906_CAD_BB_Therapy_Prior_to_MI_or_LVSD_Testing_Attachment-635975377471301600.docx

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b2-2b6)

Measure Title: CoronaryArtery Disease – Beta Blocker Therapy Prior to MI or LVSD
Date of Submission: 04/29/2016

Type of Measure:

☐ Composite – ***STOP – use composite testing form***	☐ Outcome (*including PRO-PM*)
☐ Cost/resource	☑ Process
☐ Efficiency	☐ Structure

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- **For all measures, sections 1, 2a2, 2b2, 2b3, and 2b5 must be completed.**
- **For outcome and resource use measures**, section **2b4** also must be completed.
- If specified for **multiple data sources/sets of specifications** (e.g., claims and EHRs), section **2b6** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b2-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 20 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise.

2b2. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

2b3. Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; [12]

AND

If patient preference (e.g., informed decision making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, 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). [13]

2b4. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors that influence the measured outcome (but not factors related to disparities in care or the quality of care) and are present at start of care; [14,15] and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b5. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** [16] **differences in performance**;

OR

there is evidence of overall less-than-optimal performance.

2b6. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements 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 of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality.

12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.
13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions.
15. 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, or gender (e.g., poorer treatment outcomes of African American men with prostate cancer or inequalities in treatment for CVD risk factors between men and women). It is preferable to stratify measures by race and socioeconomic status rather than to adjust out the differences.

16. 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 percent v. 75 percent) 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 less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: **(*must be consistent with data sources entered in S.23*)**	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☐ administrative claims	☐ administrative claims
☐ clinical database/registry	☐ clinical database/registry
☑ abstracted from electronic health record	☑ abstracted from electronic health record
☑ eMeasure (HQMF) implemented in EHRs	☑ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

Data 1 (EHR - Validity Against the Gold Standard)

The data source is EHR data.

Bonnie Patient Test Deck

As a supplement to the EHR reliability testing performed on this measure, a deck of patient test cases have been developed and a summary of the details has been included as part of the feasibility attachment in section 3b.3 of the measure submission form.

Data 2 (EHR – Exceptions Analysis)

The data source is EHR data.

EHR – Signal to Noise Ratio Analysis (PQRS)

The data source is EHR data from the PQRS program, provided by the Center for Medicare & Medicaid Services (CMS).

EHR –Exceptions Analysis (PQRS)

The data source is EHR data from the PQRS program, provided by the Center for Medicare & Medicaid Services (CMS).

1.3. What are the dates of the data used in testing?

Data 1 (EHR - Validity Against the Gold Standard)

The data are collected from patients sampled from 2004.

Data 2 (EHR – Exceptions Analysis)

The data are collected from patients sampled in 2009.

EHR – Signal to Noise Ratio Analysis (PQRS)

The data are for the time period January 2014 through December 2014 and cover the entire United States.

EHR –Exceptions Analysis (PQRS)

The data are for the time period January 2014 through December 2014 and cover the entire United States.

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (***must be consistent with levels entered in item S.26***)	**Measure Tested at Level of:**
☑ individual clinician	☑ individual clinician
☑ group/practice	☑ group/practice
☐ hospital/facility/agency	☐ hospital/facility/agency
☐ health plan	☐ health plan
☐ other: Click here to describe	☐ other: Click here to describe

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

Data 1 (EHR - Validity Against the Gold Standard)

The data sample came from an academic general internal medicine clinic with several years of experience using a commercial EHR. The clinic employs 40 full or part-time internal medicine physicians and provides more than 41,000 patient visits annually.

Data 2 (EHR – Exceptions Analysis)

The data sampled came from 5 physician offices using 5 different EHR systems.

EHR –Signal to Noise Ratio Analysis (PQRS)

The total number of physicians reporting on this measure, via the EHR option, in 2014, is 2,762. Of those, 473 physicians had all of the required data elements and met the minimum number of quality reporting events (10) for inclusion in the reliability analysis. For this measure, the average number of quality reporting events is 54.1 for a total 25, 605 events. The range of quality reporting events for 473 physicians included is from 836 to 10.

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

Data 1 (EHR - Validity Against the Gold Standard)

The sample consisted of approximately 134 charts for a total of 134 eligible patients. One trained investigator reviewed the 134 charts. The patients were selected using random sampling.

Data 2 (EHR – Exceptions Analysis)

The sample consisted of approximately 2,717 eligible patients.

EHR – Signal to Noise Ratio analysis (PQRS)

There were 25, 605 patients included in this reliability testing and analysis. These were the patients that were associated with physicians who had 10 or more patients eligible for this measure and remained after exceptions were removed.

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

Data 1 (EHR - Validity Against the Gold Standard)

The data sample was used for the purposes of reliability and validity testing.

Data 2 (EHR – Exceptions Analysis)

The data sample was used for the exception analysis only.

EHR – Signal to Noise Ratio analysis (PQRS)

The same data sample was used for reliability testing and exceptions analysis.

1.8. What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example, patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).

Data 1 (EHR - Validity Against the Gold Standard)

Patient-level sociodemographic (SDS) variables were not analyzed in this project.

Data 2 (EHR – Exceptions Analysis)

This was not captured as part of the testing.

EHR – Signal to Noise Ratio analysis (PQRS)

Patient-level socio-demographic (SDS) variables were not captured as part of the testing for this measure.

2a2. RELIABILITY TESTING

***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)
☑ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)
☑ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

Data 1 (EHR - Validity Against the Gold Standard)

See 2b2.2 for Validity Against the Gold Standard Results

EHR – Signal to Noise Ratio analysis (PQRS)

Reliability of the computed measure score was measured as the ratio of signal to noise. The signal in this case is the proportion of the variability in measured performance that can be explained by real differences in physician performance. Reliability at the level of the specific physician is given by:

Reliability = Variance (physician-to-physician) / [Variance (physician-to-physician) + Variance (physician-specific-error]

Reliability is the ratio of the physician-to-physician variance divided by the sum of the physician-to-physician variance plus the error variance specific to a physician. A reliability of zero implies that all the variability in a measure is

attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in physician performance.

Reliability testing was performed by using a beta-binomial model. The beta-binomial model assumes the physician performance score is a binomial random variable conditional on the physician's true value that comes from the beta distribution. The beta distribution is usually defined by two parameters, alpha and beta. Alpha and beta can be thought of as intermediate calculations to get to the needed variance estimates.

Reliability is estimated at two different points, at the minimum number of quality reporting events for the measure and at the mean number of quality reporting events per physician.

2a2.3. For each level checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

Data 1 (EHR - Validity Against the Gold Standard)

See 2b2.3 for Validity Against the Gold Standard Results

EHR – Signal to Noise Ratio analysis (PQRS)

This measure has 0.69 reliability when evaluated at the minimum level of quality reporting events and 0.92 reliability when evaluated at the average number of quality events.

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

Data 1 (EHR - Validity Against the Gold Standard)

See 2b2.4 for Validity Against the Gold Standard Results

EHR – Signal to Noise Ratio analysis (PQRS)

Reliability at the minimum level of quality reporting events is moderate. Reliability at the average number of quality events is very high.

2b2. VALIDITY TESTING

2b2.1. What level of validity testing was conducted? (*may be one or both levels*)

☑ **Critical data elements** (*data element validity must address ALL critical data elements*)

Performance measure score

☐ **Empirical validity testing**

Systematic assessment of face validity of performance measure score as an indicator of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)

2b2.2. For each level checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

Data 1 (EHR - Validity Against the Gold Standard)

Data abstracted from randomly sampled patient records were used to evaluate parallel forms reliability for the measure. Charts for abstraction were selected for patients aged 18 years and older with a visit, problem list, or medical history diagnosis of CAD.

Face Validity Assessment

Face validity of the measure score as an indicator of quality was systematically assessed as follows.

After the measure was fully specified, the expert panel was asked to rate their agreement with the following statement:

The scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good and poor quality.

Scale 1-5, where 1= Strongly Disagree; 2= Disagree 3= Neither Agree nor Disagree; 4=Agree

5= Strongly Agree

Our expert panel included 12 members. Panel members were comprised of experts from the AMA-PCPI Measure Advisory Committee. The list of expert panel members is as follows:

Amy Sanders, MD, MS

David Seidenwurm, MD

Dianne V. Jewell, PT, DPT, PhD, CCS, FAACVPR

Janet Sullivan, MD
John Easa, MD, FIPP

Joseph P. Drozda, Jr., MD, FACC

Mark Metersky, MD

Martha J. Radford, MD, FACC, FAHA

Michael O'Dell, MD, MS, MSHA, FAAFP

Richard Bankowitz, MD, MBA, FACP

Scott T. MacDonald, MD

Shannon Sims, MD, PhD

NQF Requirements of Inclusion of ICD-10 Codes:

NQF ICD-10-CM Requirement 1: Goal was to convert this measure to a new code set, fully consistent with the original intent of the measure.

NQF ICD-10-CM Requirement 2: See attachment in S.2b

NQF ICD-10-CM Requirement 3: The PCPI uses the General Equivalence Mappings (GEMs) as a first step in the identification of ICD-10 codes. We then review the ICD-10 codes to confirm their inclusion in the measure is consistent with the measure intent, making additions or deletions as needed. We have two RHIA-credentialed professionals on our staff who review all ICD-10 coding. For measures included in PQRS, the ICD-10 codes have also been reviewed and vetted by the CMS contractor. Comments received from stakeholders related to ICD-10 coding are first reviewed internally. Depending on the nature of the comment received, we also engage clinical experts to advise us as to whether a change to the specifications is warranted.

2b2.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

Data 1 (EHR - Validity Against the Gold Standard)

Of the 134 patients sampled via automated EHR review, 111 patients (82.8%) that met the numerator criteria were detected. Performance on the measure was calculated to be 90.3% through comparison of automated and manual EHR review.

Face Validity Assessment

The results of the expert panel rating of the validity statement were as follows: N = 12; Mean rating = 4.17 and 91.7% of respondents either agree or strongly agree that this measure can accurately distinguish good and poor quality.

Frequency Distribution of Ratings

1 – 1 response (Strongly Disagree)

2 – 0 responses

3 – 0 responses (Neither Agree nor Disagree)

4 – 6 responses

5 – 5 responses (Strongly Agree)

2b2.4. What is your interpretation of the results in terms of demonstrating validity? (i*.e., what do the results mean and what are the norms for the test conducted?*)

Data 1 (EHR - Validity Against the Gold Standard)

Discrepancies between performance measures based on EHR automated review alone and those based on automated review plus manual reviews were due to two types of misclassification: failure to correctly identify performance of quality measures among true, eligible patients; and failure to correctly exclude patients. Upon further analysis, the differences between automated review alone and automated plus manual reviews were 10 patients (7.5%).

Face Validity Assessment

Given that the vast majority of expert panel members agreed that the measure can accurately distinguish good and poor quality, the measure is valid, as specified.

2b3. EXCLUSIONS ANALYSIS

NA no exclusions — *skip to section 2b4*

2b3.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)
Data 1 (EHR - Validity Against the Gold Standard)

This data sample was not used to test exclusions.

Data2 (EHR – Exceptions Analysis)

Exceptions included documentation of medical reason(s), patient reason(s) and system reason(s) for not prescribing beta-blocker therapy. Exceptions were analyzed for frequency and variability across providers.

2b3.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)
Data 1 (EHR - Validity Against the Gold Standard)

This data sample was not used to test exclusions.

Data 2 (EHR – Exceptions Analysis)

Review of the 2,717 exceptions revealed that 2,292 (84.4%) exceptions were medical reasons for not prescribing beta blocker therapy, 347 (12.8%) exceptions were patient reasons and 78 (2.9%) were system reasons.

2b3.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

Exceptions are necessary to account for those situations when it is not medically appropriate to prescribe beta blocker therapy. Exceptions are discretionary and the methodology used for measure exception categories are not uniformly relevant across all measures; for this measure, there is a clear rationale to permit an exception for medical, patient or system reasons. Rather than specifying an exhaustive list of explicit medical, patient or system reasons for exception for each measure, the measure developer relies on clinicians to link the exception with a specific reason for the decision not to prescribe beta blocker therapy required by the measure.

Some have indicated concerns with exception reporting including the potential for physicians to inappropriately exclude patients to enhance their performance statistics. Research has indicated that levels of exception reporting occur infrequently and are generally valid (Doran et al., 2008), (Kmetik et al., 2011). Furthermore, exception reporting has been found to have substantial benefits: "it is precise, it increases acceptance of [pay for performance] programs by physicians, and it ameliorates perverse incentives to refuse care to "difficult" patients." (Doran et al., 2008).

Although this methodology does not require the external reporting of more detailed exception data, the measure developer recommends that physicians document the specific reasons for exception in patients' medical records for purposes of optimal patient management and audit-readiness. We also advocate for the systematic review and analysis of each physician's exceptions data to identify practice patterns and opportunities for quality improvement.

Without exceptions, the performance rate would not accurately reflect the true performance of that physician. This would result in an increase in performance failures and false negatives. The additional value of increased data collection of capturing an exception greatly outweighs the reporting burden.

References:

Doran T, Fullwood C, Reeves D, Gravelle H, Roland M. Exclusion of pay for performance targets by English Physicians. New Engl J Med. 2008; 359: 274-84.

Kmetik KS, Otoole MF, Bossley H et al. Exceptions to Outpatient Quality Measures for Coronary Artery Disease in Electronic Health Records. Ann Intern Med. 2011;154:227-234.

2b4. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b4.1. What method of controlling for differences in case mix is used?

☑ **No risk adjustment or stratification**

☐ **Statistical risk model with** Click here to enter number of factors **risk factors**

☐ **Stratification by** Click here to enter number of categories **risk categories**

☐ **Other,** Click here to enter description

2b4.2. If an outcome or resource use measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

Not applicable

2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care and not related to disparities*)

Not applicable

2b4.4a. What were the statistical results of the analyses used to select risk factors?

Not applicable

2b4.4b. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach *(describe the steps – do not just name a method; what statistical analysis was used)*

Not applicable

2b4.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)
Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
if stratified, skip to 2b4.9

Not applicable

2b4.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

Not applicable

2b4.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

Not applicable

2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

Not applicable

2b4.9. Results of Risk Stratification Analysis:

Not applicable

2b4.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)

Not applicable

***2b4.11. Optional Additional Testing for Risk Adjustment** (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods*)

Not applicable

2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b5.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

Data 1 (EHR - Validity Against the Gold Standard)

This data sample was not used to test for meaningful differences in performance across providers or practice sites.

EHR –Signal to Noise Ratio analysis (PQRS)

Measures of central tendency, variability, and dispersion were calculated.

2b5.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

Data 1 (EHR - Validity Against the Gold Standard)

This data sample was not used to test for meaningful differences in performance across providers or practice sites.

EHR – Signal to Noise Ratio analysis (PQRS)

Based on the sample of 473 included physicians, the mean performance rate is 0.57 the median performance rate is 0.53 and the mode is 0.5. The standard deviation is 0.25. The range of the performance rate is 1.00, with a minimum rate of 0 and a maximum rate of 1. The interquartile range is 0.4 (0.8 – 0.4).

2b5.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

Data 1 (EHR - Validity Against the Gold Standard)

This data sample was not used to test for meaningful differences in performance across providers or practice sites.

EHR – Signal to Noise Ratio analysis (PQRS)

The range of performance from 0.00 to 1.00 suggests there is clinically meaningful variation across physicians' performance.

2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

Note: *This criterion is directed to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***If comparability is not demonstrated, the different specifications should be submitted as separate measures.***

2b6.1. Describe the method of testing conducted to demonstrate comparability of performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

This test was not performed for this measure.

2b6.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

This test was not performed for this measure.

2b6.3. What is your interpretation of the results in terms of demonstrating comparability of performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

This test was not performed for this measure.

2b7. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b7.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (describe the steps – do not just name a method; what statistical analysis was used)

Data are not available to complete this testing.

2b7.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each)

Data are not available to complete this testing.

2b7.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias? (i.e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data)

Data are not available to complete this testing.

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without

undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. **Data Elements Generated as Byproduct of Care Processes.**

Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. **To what extent are the specified data elements available electronically in defined fields?** *(i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields)*

ALL data elements are in defined fields in electronic health records (EHRs)

3b.2. **If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources.**

3b.3. **If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL.**

Attachment **Attachment:** NQF_2906_Feasibility_Assessment_Attachment.pdf

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.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 and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.**

IF a PRO-PM, consider implications for both individuals providing PROM data (patients, service recipients, respondents) and those whose performance is being measured.

We have not identified any areas of concern or made any modifications as a result of testing and operational use of the measure in relation to data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, and other feasibility issues unless otherwise noted.

3c.2. **Describe any fees, licensing, or other requirements to use any aspect of the measure as specified** *(e.g., value/code set, risk model, programming code, algorithm)*.

The Measures, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, eg, use by health care providers in connection with their practices. Commercial uses of the Measures require a license agreement between the user and the AMA, (on behalf of the PCPI), ACC or AHA.

Limited proprietary coding is contained in the Measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals

or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Planned	Current Use (for current use provide URL)
	Public Reporting Physician Quality Reporting System (PQRS) http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/MeasuresCodes.html Payment Program Meaningful Use Stage 2 (EHR Incentive Program) http://www.cms.gov/Regulations-and-Guidance/Legislation/EHRIncentivePrograms/eCQM_Library.html

4a.1. For each CURRENT use, checked above, provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included

1. Physician Quality Reporting System (PQRS)-Sponsored by the Centers for Medicare and Medicaid Services (CMS)
Purpose: PQRS is a national reporting program that uses a combination of incentive payments and payment adjustments to promote reporting of quality information by eligible professionals (EPs). The program provides an incentive payment to practices with EPs (identified on claims by their individual National Provider Identifier [NPI] and Tax Identification Number [TIN]). Eps satisfactorily report data on quality measures for covered Physician Fee Schedule (PFS) services furnished to Medicare Part B Fee-for-Service (FFS) beneficiaries (including Railroad Retirement Board and Medicare Secondary Payer). Beginning in 2015, the program also applies a payment adjustment to EPs who do not satisfactorily report data on quality measures for covered professional services in 2013.
Source: http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/index.html CMS has implemented a phased approach to public reporting performance information on the Physician Compare Web site. CMS announced through rulemaking their plans to make all PQRS individual EP level PQRS measures available for public reporting annually, including making the 2016 PQRS individual EP level data available for public reporting on Physician Compare in late 2017.

2. Meaningful Use Stage 2 (EHR Incentive Program) – Sponsored by the Centers for Medicare and Medicaid Services (CMS)
The Medicare and Medicaid EHR Incentive Programs provide incentive payments to eligible professionals, eligible hospitals, and critical access hospitals (CAHs) as they adopt, implement, upgrade or demonstrate meaningful use of certified EHR technology. Eligibility for incentive payments for the "meaningful use" of certified EHR technology is established if all program requirements are met, including successful implementation and reporting of program measures, which include this measure, to demonstrate meaningful use of EHR technology.

4a.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)
We support the expanded use of this measure in government or other programs, including those intended for accountability or public reporting. The ACC, AHA and PCPI do not have any policies that would restrict access to the performance measure specifications or results or that would impede implementation of the measure for any application. We would welcome its implementation in emerging applications such as accountable care organizations (ACO), Medicare Advantage insurance plans or health plans selling on the new insurance marketplace.

4a.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*Credible plan includes the specific program, purpose, intended audience, and timeline for*

implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.)
As described above, it is our understanding that CMS is also planning to move towards publicly reporting physician data via Physician Compare. Also, although the measure is currently in use, we support expanded use of this measure in government or other programs, including those intended for accountability or public reporting.

4b. Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b.1. **Progress on Improvement. (Not required for initial endorsement unless available.)**
Performance results on this measure (current and over time) should be provided in 1b.2 and 1b.4. Discuss:

- **Progress (trends in performance results, number and percentage of people receiving high-quality healthcare)**
- **Geographic area and number and percentage of accountable entities and patients included**

2014 PQRS Experience Report
2014 is the most recent year for which PQRS Experience Report measure data is available. The average performance rates on Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%) over the last several years are as follows:
2010: 71.4%
2011: 82.1%
2012: 69.9%
2013: 74.2%
2014: 79.3%

It is important to note that PQRS has been and remains a voluntary reporting program. In the early years of the PQRS program, participants received an incentive for satisfactorily reporting. However, beginning in 2015, the program imposes payment penalties for non-participants based on 2013 performance. 62% of eligible professionals participated using any reporting option in 2014. As a result, performance rates may not be nationally representative.

Reference: Center for Medicare and Medicaid Services. 2014 Reporting Experience Including Trends (2007-2014). Available at: https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/pqrs/analysisandpayment.html

4b.2. **If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.**
While we create measures with an ultimate goal of improving the quality of care, measurement is a mechanism to drive improvement but does not equate with improvement. Measurement can help identify opportunities for improvement with actual improvement requiring making changes to health care processes and structure. In order to promote improvement, quality measurement systems need to provide feedback to front-line clinical staff in as close to real time as possible and at the point of care whenever possible. (1)

1. Conway PH, Mostashari F, Clancy C. The future of quality measurement for improvement and accountability. JAMA. 2013 Jun 5;309(21):2215-6.

4c. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4c.1. **Were any unintended negative consequences to individuals or populations identified during testing; OR has evidence of unintended negative consequences to individuals or populations been reported since implementation? If so, identify the negative unintended consequences and describe how benefits outweigh them or actions taken to mitigate them.**
We are not aware of any unintended consequences at this time, but we take unintended consequences very seriously and therefore continuously monitor to identify actions that can be taken to mitigate them.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures
Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.
Yes

5.1a. **List of related or competing measures (selected from NQF-endorsed measures)**
0070 : Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)
0071 : Persistence of Beta-Blocker Treatment After a Heart Attack
0083 : Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)
2908 : Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)

5.1b. **If related or competing measures are not NQF endorsed please indicate measure title and steward.**

5a. Harmonization

The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. **If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):**
Are the measure specifications completely harmonized?
No

5a.2. **If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.**
eMeasure 2906 addresses a patient population of patients with CAD and either a recent prior MI or LVSD. This patient population is also covered in part by the following NQF-endorsed measures: NQF 0071: Persistence of Beta-Blocker Treatment After a Heart Attack and NQF 0083 and 2908: Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD). The specifications are harmonized to the extent possible. As a result, the denominator specifications for the measures differ where needed based on the differing patient populations. Measure 0070 is the registry version of this eMeasure and is completely harmonized.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. **If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):**
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required

attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment:

Contact Information

Co.1 **Measure Steward (Intellectual Property Owner):** PCPI Foundation
Co.2 **Point of Contact:** Samantha, Tierney, Samantha.Tierney@ama-assn.org, 312-464-5524-
Co.3 **Measure Developer if different from Measure Steward:** PCPI Foundation
Co.4 **Point of Contact:** Samantha, Tierney, Samantha.Tierney@ama-assn.org, 312-464-5524-

Additional Information

Ad.1 **Workgroup/Expert Panel involved in measure development**
Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
Work Group members:
Joseph Drozda, MD, FACC (Co-Chair) (cardiology; methodology)
Joseph V. Messer, MD, MACC, FAHA (Co-Chair) (cardiology)
John Spertus, MD, FACC, FAHA (Co-Chair) (cardiology)
Bruce Abramowitz, MD, FACC (interventional cardiology; measure implementation)
Karen Alexander, MD, FACC (cardiology; geriatrics)
Craig T. Beam, CRE (patient representative)
Robert O. Bonow, MD, MACC, FAHA, FACP (cardiology)
Jill S. Burkiewicz, PharmD, BCPS (pharmacy)
Michael Crouch, MD, MSPH (family medicine)
David C. Goff, Jr., MD, PhD, FAHA, FACP (internal medicine)
Richard Hellman, MD, FACP, FACE (endocrinology)
Thomas James, III, FACP, FAAP (health plan representative)
Marjorie L. King, MD, FACC, MAACVPR (cardiology; cardiac rehabilitation)
Edison A. Machado, Jr., MD, MBA (measure implementation)
Eduardo Ortiz, MD, MPH (internal medicine; guideline development)
Michael O'Toole, MD, FACC (cardiology; electrophysiology; measure implementation)
Stephen D. Persell, MD, MPH (internal medicine; measure implementation)
Jesse M. Pines, MD, MBA, MSCE, FAAEM (emergency medicine)
Frank J. Rybicki, MD, PhD (radiology)
Lawrence B. Sadwin (patient representative)
Joanna D. Sikkema, MSN, ANP-BC, FAHA (cardiology)
Peter K. Smith, MD (thoracic surgery)
Patrick J. Torcson, MD, FACP, MMM (hospital medicine)
John B. Wong MD, FACP (internal medicine)

ACCF, AHA, and PCPI measures are developed through cross-specialty, multi-disciplinary work groups. All medical specialties and other health care professional disciplines participating in patient care for the clinical condition or topic under study are invited to participate as equal contributors to the measure development process. In addition, the ACCF, AHA and PCPI strive to include on their work groups individuals representing the perspectives of patients, consumers, private health plans, and employers. This broad-based approach to measure development ensures buy-in on the measures from all stakeholders and minimizes bias toward any individual specialty or stakeholder group. All work groups have at least two co-chairs who have relevant clinical and/or measure development expertise and who are responsible for ensuring that consensus is achieved and that all perspectives are voiced.

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.2 **Year the measure was first released:** 2003
Ad.3 **Month and Year of most recent revision:** 06, 2015
Ad.4 **What is your frequency for review/update of this measure?** Supporting guidelines and specifications for this measure are reviewed on an annual basis.
Ad.5 **When is the next scheduled review/update for this measure?** 06, 2016

Ad.6 **Copyright statement:** Copyright 2015 American College of Cardiology, American Heart Association and American Medical Association. All Rights Reserved.

Ad.7 **Disclaimers:** The Measures are not clinical guidelines, do not establish a standard of medical care, and have not been tested for all potential applications.

The Measures, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, eg, use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the Measures for commercial gain, or incorporation of the Measures into a product or service that is sold, licensed or distributed for commercial gain.

Commercial uses of the Measures require a license agreement between the user and the PCPI(R) Foundation (PCPI[R]), the American Medical Association (AMA), American College of Cardiology (ACC) or the American Heart Association (AHA). Neither the AMA, nor ACC, nor AHA, nor the AMA-convened Physician Consortium for Performance Improvement(R) (AMA-PCPI), now known as the PCPI, nor their members shall be responsible for any use of the Measures.

AMA and PCPI encourage use of the Measures by other health care professionals, where appropriate.

THE MEASURES AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

Limited proprietary coding is contained in the Measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. The AMA, ACC, AHA, the PCPI and its members and former members of the AMA-PCPI disclaim all liability for use or accuracy of any Current Procedural Terminology (CPT[R]) or other coding contained in the specifications.

CPT(R) contained in the Measure specifications is copyright 2004-2015 American Medical Association. LOINC(R) is copyright 2004-2015 Regenstrief Institute, Inc. This material contains SNOMED CLINICAL TERMS (SNOMED CT[R]) copyright 2004-2015 International Health Terminology Standards Development Organisation (IHTSDO). ICD-10 is copyright 2015 World Health Organization. All Rights Reserved.

Ad.8 **Additional Information/Comments:**

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 0290

De.2. **Measure Title:** Median Time to Transfer to Another Facility for Acute Coronary Intervention

Co.1.1. **Measure Steward:** Centers for Medicare & Medicaid Services

De.3. **Brief Description of Measure:** This measure calculates the median time from emergency department (ED) arrival to time of transfer to another facility for acute coronary intervention (ACI) for ST-segment myocardial infarction (STEMI) patients that require a percutaneous coronary intervention (PCI). The measure is calculated using chart-abstracted data, on a rolling quarterly basis, and is publically reported, in aggregate, for one calendar year. The measure has been publically reported, annually by CMS as a component of its Hospital Outpatient Quality Reporting (HOQR) Program since 2008.

1b.1. **Developer Rationale:** The early use of primary angioplasty in patients with STEMI results in a significant reduction in mortality and morbidity. The earlier PCI is provided, the more effective it is (Rathore, 2009). National guidelines recommend the prompt initiation of PCI in patients presenting with STEMI infarction (O'Gara, 2013). Current recommendations support a door-to-balloon time of 120 minutes or less in patients that need to be transferred from a non-PCI capable hospital to a PCI-capable hospital (O'Gara, 2013). Patients transferred for primary PCI (pPCI) rarely meet recommended guidelines for door-to-balloon time (Dauerman et. al., 2015). Elevated transfer time is a significant predictor of delay in the initiation of PCI (Dauerman et. al., 2015). Decreasing transfer time in STEMI patients requiring an acute coronary intervention has the potential to lead to reduced door-to balloon-time.

REFERENCES:

1) Dauerman HL, Bates ER, Kontos MC, Li S, Garvey JL, Henry TD, Manoukian SV, Roe MT. Nationwide analysis of patients With ST-segment-elevation myocardial infarction transferred for primary percutaneous intervention: findings from the american heart association mission: lifeline program. Circ Cardiovasc Interv. 2015:8(5):pii: e002450. doi: 10.1161/CIRCINTERVENTIONS.114.002450.

2) O´Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA, Granger CB, Krumholz HM, Linderbaum JA, Morrow DA, Newby LK, Ornato JP, Ou N, Radford MJ, Tamis-Holland JE, Tommaso CL, Tracy CM, Woo YJ, Zhao DX. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013;61(4):e78-140.

3) Rathore SS, Curtis JP, Chen J, et. al. Association of door-to-balloon time and mortality in patients admitted to hospital with ST elevation myocardial infarction: national cohort study. BMJ?: British Medical Journal. 2009;338:b1807. doi:10.1136/bmj.b1807.

S.4. **Numerator Statement:** This measure is reported as a continuous variable statement: time (in minutes) from ED arrival to transfer to another facility for ACI.

The numerator includes patients with AMI and ST-segment elevation on the ECG performed closest to ED arrival who are transferred from the ED to a short-term general hospital for inpatient care, or to a Federal healthcare facility specifically for ACI.

S.7. **Denominator Statement:** Time (in minutes) from ED arrival to transfer to another facility for ACI.

S.10. **Denominator Exclusions:** Patients are excluded from this measure if they are under 18 years of age, did not have an initial ECG interpretation, received fibrinolytic therapy while in the ED, or were transferred for reasons other than ACI.

De.1. **Measure Type:** Process

S.23. **Data Source:** Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

S.26. **Level of Analysis:** Facility, Population : National

IF Endorsement Maintenance – Original Endorsement Date: Nov 15, 2007 **Most Recent Endorsement Date:** Jan 18, 2012

Maintenance of Endorsement-- Preliminary Analysis

To maintain NQF endorsement endorsed measures are evaluated periodically to ensure that the measures still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence
Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

1a. Evidence. The evidence requirements for a *process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Summary of prior review in 2012:

- The steps between the measure focus and the health outcome include decreasing transfer time in STEMI patients requiring an acute coronary intervention from a non-PCI-capable hospital to a PCI-capable hospital has the potential to lead to reduced door-to-balloon time, which leads to a decrease in mortality.
- The developer provided one clinical guideline from the 2004 ACC/AHA Guidelines for the Management of Patients with ST-Elevation Myocardial Infarction:
    - if PCI is chosen, the delay from patient contact with the healthcare system (typically, arrival at the ED or contact with paramedics) to balloon inflation should be less than 90 minutes. **Level of Evidence**: **B** (Level B: Other evidence)

Changes to evidence from last review

☐ **The developer attests that there have been no changes in the evidence since the measure was last evaluated.**

☒ **The developer provided updated evidence for this measure:**

Updates:

- The developer provided an additional clinical guideline from the 2013 ACCF/AHA Guideline for the Management of ST-Elevation Myocardial Infarction that provides two recommendations for the transfer of patients who require primary PCI (pPCI), from a non-PCI-capable hospital to a PCI-capable hospital, in cases where pPCI can be performed within 120 minutes of first medical contact (FMC):
    - Immediate transfer to a PCI-capable hospital for primary PCI is the recommended triage strategy for patients with STEMI who initially arrive at or are transported to a non-PCI capable hospital, with an FMC-to-device time system goal of 120 minutes or less. (**Class I, Level of Evidence**: **B**)
    - Immediate transfer to a PCI-capable hospital for coronary angiography is recommended for suitable patients with STEMI who develop cardiogenic shock or acute severe HF, irrespective of the time delay from MI onset. (**Class I, Evidence: B**)
- The developer provided a systematic review of the body of the evidence supporting the timely transfer of STEMI patients requiring a PCI. The details of the Quality, Quantity, and Consistency of the evidence provided was associated with the guideline (single randomized trial or non-randomized studies).
- The developer identified five new studies that were published since the systematic review of the body of evidence (2002-2012).

Exception to evidence : N/A

Guidance from the Evidence Algorithm: Process measure/Systematic review (Box 3)→ QQC associated w/guideline presented (Box 4)→Systematic review concluded: Quantity: Low; Quality: Moderate; Consistency: Moderate (Box

5b)→Moderate

Questions for the Committee:

- *Is the Committee willing to accept the prior evaluation? The updated guideline supports the measure focus and has a strong level of evidence.*

Preliminary rating for evidence: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities
Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- The developer provided the following facility-level performance rates from Hospital Compare from the April 2010 – March 2015 data collection period:

	2010-11	2011-12	2012-13	2013-14	2014-15	2010-15 Change in Minutes
Facilities	421	400	405	409	425	--
Denominator	8,008	7,621	7,822	7,678	8,166	--
Minimum time	26	20	14	21	20	-6
1st PCTL	26	22	24	25	26	--
5th PCTL	32	31	30	30	30	-2
10th PCTL	36	35	34	34	33	-3
25th PCTL	45	44	42	41	42	-3
Median	56	54	54	54	54	-2
75th PCTL	70	68	68	71	69	-1
90th PCTL	91	90	89	87	94	3
95th PCTL	104	120	110	108	113	9
99th PCTL	210	229	152	166	206	-4
Maximum time	245	542	307	288	474	229

- The developer noted that the performance scores represented facilities whose denominator counts met minimum case count requirements during the April 2010-March 2015 data collection periods but does not state the minimum case count.

Disparities:

- The developer analyzed the relationship of patient and facility characteristics on time to transfer from the ED to another facility for ACI using 2014 data submitted to the Clinical Data Warehouse (CDW).
- The results indicated that age, race, sex, and facility characteristics were variables related to timely transfer for ACI:
  - Patients aged 40-60 had significantly lower time to transfer than patients aged 18-30.
  - African-American and Asian patients had a significantly longer time to transfer compared to White patients.
  - Hispanic patients had a significantly longer time to transfer compared to non-Hispanic patients.
  - Female patients had a significantly longer time to transfer compared to male patients.
  - Patents transferred from facilities with fewer than 50 beds had a significantly lower time to transfer than patients transferred from facilities with 101-500+ beds.
  - Patients transferred from urban hospitals and teaching facilities had significantly lower time to transfer

than patients transferred from rural hospitals and non-teaching facilities.

Questions for the Committee:

- *Does the data presented demonstrate that there continues to be a quality problem and variation with STEMI patients that arrive at the ED and require transfer to another facility for acute coronary intervention? Is there an opportunity for improvement?*
- *Is a national performance measure still warranted?*
- *Are you aware of evidence that additional disparities exist in this area of healthcare?*

Preliminary rating for opportunity for improvement: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence to Support Measure Focus

Comments: **This is a process measure. The evidence is from a clinical practice guideline, 2013, and is Class I, Level B. The evidence is directly related to the measure. The guideline included an evidence review and 5 new studies since the last review of the measure. The new studies continue to emphasize the time-sensitive nature of rapid time to treatment.

**This is a process measure that has been associated with outcomes; a longer time to transfer is a predictor of delay to receiving PCI and delay in receiving PCI has been associated with worse patient outcomes.

1b. Performance Gap

Comments: **The median time to transfer at the facility level showed no improvement from 2011-2015 (54 minutes, IQR 42-69 minutes, 10th-90th percentiles: 33-94 minutes), demonstrating significant facility-level variation. Subgroup analyses found blacks, Hispanics, and Asians had significantly longer times, as did rural and non-teaching hospitals.

**there appears to be no change in the times measured over time

disparities were noted and exist"

**The performance gap has remained the same but the issue gets into understanding the prevelance and moving well beyond that time frames that the patients that are yours

**The performance gap is concerning, and data on disparities are compelling.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

2a1. Reliability Specifications

Maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented.

Data source(s): Administrative claims, Electronic Health Record, Paper Medical Records

Specifications:

- This is a facility-level measure.
- This measure is a continuous measure; therefore, the numerator and denominator are the same: Time (in minutes) from ED arrival to transfer to another facility for acute coronary intervention (ACI). The measure population includes:
  - Patients with AMI and ST-segment elevation on the ECG performed closest to ED arrival (not more than 1 hour prior to arrival)
  - Patients who do not receive fibrinolytic therapy in the ED
  - Documentation the patient was transferred from facility's ED to another facility specifically for ACI.
- Denominator exclusions include:
  - Patients under 18 years of age
  - Patients with no ST-elevation on the interpretation of the 12-lead ECG performed closest to ED arrival, no interpretation or report available, or UTD from documentation.

- Patients with documentation that fibrinolytic therapy was initiated at the ED
- Patients admitted to observation status prior to transfer and patients transferred for reasons other than ACI or specific reason for transfer was unable to be determined from documentation.

- Discharge codes, evaluation and management (E/M) codes, and ICD-10 codes included in attachment NQF_0290_MeasureCodeSet.xlsx.
- The calculation algorithm is included.
- Instructions for sampling and guidance on minimum sample size are provided.
- Details for handling missing data are provided.
- An electronic data collection tool is available from vendors. Free data collection tools are also available on qualitynet.org.

Questions for the Committee :

- *Are all the data elements clearly defined? Are all appropriate codes included?*
- *Is the logic or calculation algorithm clear?*
- *Is it likely this measure is consistently implemented?*

2a2. Reliability Testing attachment
Maintenance measures – less emphasis if no new testing data provided

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers.

For maintenance measures, summarize the reliability testing from the prior review:

- At the time of the last review the measure was undergoing validation through the CMS Clinical Abstraction Center.

Describe any updates to testing: measure score level testing conducted – see below

SUMMARY OF TESTING

Reliability testing level ☒ **Measure score** ☐ **Data element** ☐ **Both**

Reliability testing performed with the data source and level of analysis indicated for this measure ☒ **Yes** ☐ **No**

Method(s) of reliability testing:

- The dataset included 1,902 facilities that submitted 64,827 cases to Hospital Compare from April 1, 2014-December 31, 2014. Of those, a total of **13,195** cases remained in the denominator after the exclusions were applied. Facilities with fewer than **11** cases were omitted in accordance with Hospital Compare's minimum case count criteria.
- The developers used a beta-binomial model to assess the signal-to-noise ratio to test the reliability of the measure score. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in performance. The higher the reliability score, the greater is the confidence with which one can distinguish the performance of one facility from another. This is an appropriate test for measure score reliability. A reliability of 0.70 is generally considered a minimum threshold for reliability.

Results of reliability testing:

- The developer stated that the reliability score was 0.78. It is not clear if this was the reliability score for the facilities meeting the minimum case count (11) or overall reliability.

Guidance from the Reliability Algorithm : Precise specifications (Box 1)→Empirical reliability testing (Box 2) →Computed performance scores for measured entities (Box 4) →Appropriate method used (Box 5) →Moderate

Questions for the Committee:

- *Is the test sample adequate to generalize for widespread implementation?*

- *Do the results demonstrate sufficient reliability so that differences in facility performance can be identified?*

Preliminary rating for reliability: ☒ **High** ☐ **Moderate** ☐ **Low** ☐ **Insufficient**

2b. Validity
Maintenance measures – less emphasis if no new testing data provided

2b1. Validity: Specifications

2b1. Validity Specifications. This section should determine if the measure specifications are consistent with the evidence.

Specifications consistent with evidence in 1a. ☒ **Yes** ☐ **Somewhat** ☐ **No**
Specification not completely consistent with evidence

Question for the Committee:
- *Are the specifications consistent with the evidence?*

2b2. Validity testing

2b2. Validity Testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

For maintenance measures, summarize the validity testing from the prior review:
- At the time of the last review the measure was undergoing validation through the CMS Clinical Abstraction Center.

Describe any updates to validity testing: data element testing – see below

SUMMARY OF TESTING
Validity testing level ☐ **Measure score** ☒ **Data element testing against a gold standard** ☐ **Both**

Validity testing method:
- The dataset used for patient-level data element testing included a sample of 462 cases submitted to the Clinical Data Warehouse (CDW) from 65 facilities from April 1, 2014-March 31, 2015. After exclusions, 86 cases remained in the denominator sample.
- Data element validity was conducted by assessing the level of agreement between facility abstraction and auditor (CDAC) abstraction (gold standard) and calculating a kappa statistic or Pearson correlation coefficient.
  - Kappa values range between 0 and 1.0 and are interpreted as degree of agreement beyond chance. By convention, a kappa > .70 is considered acceptable.
    - 0 No better than chance
    - 0.01-0.20 Slight
    - 0.21-0.40 Fair
    - 0.41-0.60 Moderate
    - 0.61-0.80 Substantial
    - 0.81-1.0 Almost perfect
  - P-values estimate the statistical significance associated with the test statistics. P-values of less than 0.001 suggest high levels of statistical significance and that the degree of agreement was not due to chance.

Validity testing results:
- The kappa and Pearson's correlation coefficient for 12 critical data elements ranged from 0.39 to 1.0.

Data Element	Test Statistic (p-value)
E/M Code [a]	1.00 (<0.001)
Discharge Code [a]	1.00 (<0.001)
Patient Age on Outpatient Encounter Date [c]	-

ICD-9-CM Principal Diagnosis Code [a]	1.00 (<0.001)
Initial ECG Interpretation [a]	**0.63** (<0.001)
Fibrinolytic Administration [a]	0.93 (<0.001)
Transfer for Acute Coronary Intervention [a]	**0.47** (<0.001)
Discharge Date [b]	1.00 (<0.001)
Discharge Time [b]	1.00 (<0.001)
Arrival Time [b]	0.99 (<0.001)
Measurement Value [c]	-
Reason for Not Administering Fibrinolytic Therapy [a]	**0.39** (<0.001)

a. The test statistic to assess validity for this data element is a Kappa score.
b. The test statistic to assess validity for this data element is a Pearson's correlation.
c. This data element is a calculated value, not an abstracted value.

Questions for the Committee:

- *Do the validity testing results demonstrate that all of the measure elements are correct?*
- *Is the test sample adequate to generalize for continued widespread implementation?*

2b3-2b7. Threats to Validity

2b3. Exclusions:

- The developer examined overall frequencies and proportions of exclusions/exceptions for 1,902 facilities that submitted eligible cases in 2014. The sampled population included 64,827 cases where a patient, age 18 years or older, presented with an AMI to the ED.

Overall Occurrence and Distribution across Facilities for Measure Exclusions and Exceptions

Data Element	Denominator Exclusion or Numerator Exception?		Overall Occurrence		Distribution Across Facilities (%)		
	Denominator Exclusion	Numerator Exception	N	%	25th	50th	75th
Initial ECG Interpretation	X		38,573	59.5	44.1	60.0	76.7
Fibrinolytic Administration	X		4,600	7.1	0.0	0.0	9.1
Transfer for Acute Coronary Intervention	X		3,386	5.2	0.0	0.0	6.9
Discharge Date*		X	25	0.0	0.0	0.0	0.0
Discharge Time*		X	104	0.2	0.0	0.0	0.0
Arrival Time*		X	7	0.0	0.0	0.0	0.0
Reason for Not Administering Fibrinolytic Therapy*		X	4,937	7.6	0.0	0.0	27.4
Total Denominator Exclusions	3 exclusions	-	46,559	71.8	58.3	77.8	100.0
Total Numerator Exceptions	-	4 exceptions	5,073	7.8	0.0	0.0	9.8
Total Removed from the Denominator or Numerator	7 exceptions and exclusions		51,632	79.6	68.7	89.2	100.0

*If these data elements are to "UTD", the case is not included in the measure numerator but remains in the effective sample.

Questions for the Committee:

- *Are the exclusions consistent with the evidence?*
- *Does the large number of denominator exclusions (>70%)represent a significant threat to validity of the measure*

results?

- *Are any patients or patient groups inappropriately excluded from the measure?*
- *Are the exclusions/exceptions of sufficient frequency and variation across providers to be needed (and outweigh the data collection burden)?*

2b4. Risk adjustment: **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

2b5. Meaningful difference *(can statistically significant and clinically/practically meaningful differences in performance measure scores can be identified)*:

- The developer provided the distribution of facility performance scores for 425 facilities from the April 2014-March 2015 data collection period.

Distribution of Facility Performance Scores								
Mean	**Std. Dev.**	**Min.**	**10th Percent**	**Lower Quartile**	**Median**	**Upper Quartile**	**90th Percent**	**Max.**
62.1	38.5	20	33	42	54	69	94	474

Question for the Committee:

- *Does this measure identify statistically and meaningful differences in performance among facilities?*

2b6. Comparability of data sources/methods:

- Measure is not specified for more than one data source; comparability of data sources is not needed.

2b7. Missing Data

- Cases with a value of "UTD" are included in the denominator but excepted from the numerator. The developer described the frequency and distribution of missing data in Section 2b3.3.

Guidance from Validity Algorithm: Specifications consistent with evidence (Box 1) →Potential threats to validity assessed (Box 2) →Empirical validity testing (Box 3)→Patient-level data element validity testing conducted; approximately 20% of cases remained in denominator sample after exclusions (Box 10) →Element data elements abstracted and compared against the gold standard and kappa/Pearson's correlation coefficient calculated (Box 11)→Moderate (highest eligible rating is MODERATE)

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2d)

2a1. & 2b1. Specifications

Comments: **This is a continuous measure of time to transfer to a PCI-capable hospital. Data elements are clearly specified. The algorithm for calculation is provided and seems clear.

**Reliability specifications are very good. Ranking: High.

**Specifications seem to be consistent with the evidence.

**Validity specifications are very good. Ranking: High.

2a2. Reliability Testing

Comments: **Reliability testing produced an intraclass correlation coefficient of 0.78. Testing was completed on 1902 facilities and 13,195 patients after removing cases for defined exclusions. Testing was done at the performance measure level with signal to noise analysis. Facilities with fewer than 11 cases were omitted from reliability testing – unclear how the cut-point was chosen.

**Reliability testing is very good. Ranking: High.

2b2. Validity Testing

Comments: **The same population was used for both reliability and validity testing. Validity testing was done on individual data elements. Of the 12 data elements, two had kappas < 0.5 (0.47 & 0.39), one had a kappa of 0.63 and the others were 0.99 or 1.00. Face validity was done using 5 stakeholders. All five either agreed or strongly agreed that the measure captures the correct population and that the measure assesses timely transfer.

**Validity testing is not as good. Ranking: Moderate. I also do not fully understand the threats to validity data the develop provided and will as them about this during our face to face meeting.

2b3. Exclusions Analysis

2b4. Risk Adjustment/Stratification for Outcome or Resource Use Measures

2b5. Identification of Statistically Significant & Meaningful Differences In Performance

2b6. Comparability of Performance Scores When More Than One Set of Specifications

2b7. Missing Data Analysis and Minimizing Bias

Comments: **Stakeholders generally agreed on the 3 exclusion criteria, although there was considerable facility-level variation in exclusions (IQR 44.1%-76.7%). Exclusions due to ECG interpretation had a kappa of 0.63 on validity testing. 59.5% of potential cases were excluded based on this exclusion criteria. Only cases with STEMI on ECG closest to arrival are included. The other 2 exclusion criteria (fibrinolytic administration and cases where the patient was admitted for observation or transfer for reasons other than for acute coronary intervention) had much less facility-level variation in use and accounted for far fewer cases. There was no information pertaining to bias for excluding certain groups more or less than others.

This is not risk-adjusted.

There are significant differences in time to transfer at the facility level (range: 20-474 minutes; mean 62.1, SD 38.5 minutes).

There were few cases with missing data, as the system is designed to reject cases with missing data.

**"See above for threats to validity.

Risk adjustment: not applicable.

Comparability of performance measures: NA.

Missing data/no response: no.

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- Some data elements are in defined fields in electronic sources.
- The developer noted that challenges in interpreting and operationalizing the current measure's algorithm make it difficult to respecify this measure for an EHR reporting program since some data elements of the measure currently rely on logic and inferences that abstractors have been trained to interpret.
- Data collection burden due to manual abstraction to collect the data.
- There are no fees, licensure, or other requirements necessary to use this measure.
- The developer stated that the majority of a five member expert panel agreed practical aspects of reporting this chart-abstracted measure do not place undue burden on facilities that collect the data though the costs of collecting the data and reporting this measure were not provided.
    - Those members of the expert panel that did not agree specifically that requiring details about the logistics of a patient transfer as well as the documentation of contraindications may pose a burden on facilities as this is not something that is normally documented in ED workflow.

Questions for the Committee:

- *Are the required data elements routinely generated and used during care delivery?*
- *Are the required data elements available in electronic form, e.g., EHR or other electronic sources?*

Preliminary rating for feasibility: ☐ **High** ☐ **Moderate** ☒ **Low** ☐ **Insufficient**

Rationale: Chart abstraction with little likelihood of electronic specifications anytime soon.

Committee pre-evaluation comments

Criteria 3: Feasibility

3a. Byproduct of Care Processes

3b. Electronic Sources

3c. Data Collection Strategy

Comments: **The measure is chart abstracted. Some but not all data elements are easily abstracted electronically.

**Feasibility rating is at low but I think could be elevated to moderate if the data elements needed for this measure are added to EHR or a registry.

Criterion 4: Usability and Use
Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact /improvement and unintended consequences

4. Usability and Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

Current uses of the measure
Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No**

Accountability program details:

- CMS HOQR Program: The HOQR Program is a pay for quality data reporting program implemented by CMS for outpatient hospital services. Hospital quality of care information gathered through the HOQR Program is publicly available on the Hospital Compare website.

Improvement results:

- The developer provided a summary of statistics of performance scores from the April 2010 – March 2015 data collection periods. The developer also stated that these cases reflect only a subset of the patients eligible for the measure. Dependent upon the facility's total case count, the facility may report all cases or a sample of cases; thus, the number of patients receiving high-quality healthcare as performance on the measure improves is larger than the number of cases captured by the measure.

Unexpected findings (positive or negative) during implementation:

- There has been wide variation in facility performance from the inception of public reporting through March 2015. While median performance is improving, there is an ongoing opportunity for improvement in facility performance. The maximum measure values are variable across data collection periods, suggesting there may be a benefit in artificially censoring outliers. Artificially censoring outlier facilities would prevent skewing median and mean measurement values without decreasing facility sample size or awarding poor performance.

Potential harms:

- The developer has not identified any unintended consequences.

***Questions for the Committee**:*

- *How can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Does the large number of exclusions limit the usability of this measure?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

Preliminary rating for usability and use: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 4: Usability and Use

4a. Accountability and Transparency
4b. Improvement
4c. Unintended Consequences
Comments: ** The measure is publicly reported through CMS Hospital Outpatient Quality Reporting. Benchmarking is provided. Extremely long times that are outliers may skew results if not censored.
**I would say moderate as this measure has been in use, but one question I will ask the developer is on data collected to date regarding the impact of this measure on processes of care.

Criterion 5: Related and Competing Measures

Related or competing measures:

- 0163 : Primary PCI received within 90 minutes of hospital arrival
- 0288 : Fibrinolytic Therapy Received Within 30 Minutes of ED Arrival

Harmonization:

- NQF #0163 #0163 is included in the Hospital Inpatient Quality Reporting (HIQR) Program as an electronically specified clinical quality measure (eCQM) and focuses on the timely initiation of PCI for a patient who arrives at a PCI-capable hospital.
- #0288 focuses on the timely administration of fibrinolytic therapy.

Pre-meeting public and member comments

-

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0290
Measure Title: Median Time to Transfer to Another Facility for Acute Coronary Intervention
IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: 4/29/2016

Instructions

- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- Respond to all questions as instructed with answers immediately following the question. All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 10 pages (*incudes questions/instructions*; minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus
The measure focus is evidence-based, demonstrated as follows:

- Health outcome: [3] a rationale supports the relationship of the health outcome to processes or structures of care. Applies to patient-reported outcomes (PRO), including health-related quality of life/functional status, symptom/symptom burden, experience with care, health-related behavior.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- Efficiency: [6] evidence not required for the resource use component.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.
4. The preferred systems for grading the evidence are the U.S. Preventive Services Task Force (USPSTF) grading definitions and methods, or Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.
5. 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 multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.
6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)
Outcome
☐ Health outcome: Click here to name the health outcome
☐Patient-reported outcome (PRO): Click here to name the PRO

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors

☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☒ Process: This measure calculates the median time from emergency department arrival to time of transfer to another facility for an acute coronary intervention for ST-segment elevation myocardial infarction (STEMI) patients that require a percutaneous coronary intervention (PCI).

☐ Structure: Click here to name the structure

☐ Other:

HEALTH OUTCOME/PRO PERFORMANCE MEASURE *If not a health outcome or PRO, skip to 1a.3*

1a.2. Briefly state or diagram the path between the health outcome (or PRO) and the healthcare structures, processes, interventions, or services that influence it.

This measure is not a health outcome/PRO performance measure.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process, intervention, or service (*i.e., influence on outcome/PRO*).

This measure is not a health outcome/PRO performance measure.

Note: For health outcome/PRO performance measures, no further information is required; however, you may provide evidence for any of the structures, processes, interventions, or service identified above.

INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

1a.3. Briefly state or diagram the path between structure, process, intermediate outcome, and health outcomes. Include all the steps between the measure focus and the health outcome.

The early use of primary angioplasty in patients with ST-segment myocardial infarction (STEMI) results in a significant reduction in mortality and morbidity. The earlier primary coronary intervention is provided, the more effective it is (Brodie, 1998 and DeLuca, 2004). National guidelines recommend the prompt initiation of percutaneous coronary intervention (PCI) in patients presenting with ST-segment elevation myocardial infarction (Antman, 2004). Patients transferred for primary PCI rarely meet recommended guidelines for door-to-balloon time (Nallamothu, 2005). Times to treatment in transfer patients undergoing primary PCI may influence the use of PCI as an intervention (Nallamothu, 2005). Current recommendations support a door-to-balloon time of 90 minutes or less (Krumholz, 2008).

The early use of primary angioplasty in patients with STEMI results in a significant reduction in mortality and morbidity. The earlier primary percutaneous coronary intervention (PCI) is provided, the more effective it is (Rathore, 2009). National guidelines recommend the prompt initiation of PCI in patients presenting with STEMI (O'Gara, 2013). Current recommendations support a door-to-balloon time of 120 minutes or less in patients that need to be transferred from a non-PCI-capable hospital to a PCI-capable hospital (O'Gara, 2013). In a nationwide study of 14,518 patients transferred from non-PCI-capable hospitals to PCI-capable hospitals, more than one-third of patients failed to meet recommended guidelines for door-to-balloon time (Dauerman et al, 2015).

Studies have shown that delays in treatment are associated with an increased risk-adjusted in-hospital mortality in a continuous, non-linear fashion (Rathore, 2009). A reduction in door-to-balloon time from 150 minutes to 120 minutes was associated with 1.4% lower mortality; the risk of mortality continued to decrease with reductions in door-to-balloon time (Rathore, 2009). Because elevated transfer time has been shown to be a significant predictor of delay in the initiation of PCI (Dauerman et al, 2015), decreasing transfer time in STEMI patients requiring an acute coronary intervention has the potential to lead to reduced door-to balloon time.

REFERENCES:

1) Dauerman HL, Bates ER, Kontos MC, Li S, Garvey JL, Henry TD, Manoukian SV, Roe MT. Nationwide analysis of patients With ST-segment-elevation myocardial infarction transferred for primary percutaneous intervention: findings from the american heart association mission: lifeline program. Circ Cardiovasc Interv. 2015:8(5):pii: e002450. doi: 10.1161/CIRCINTERVENTIONS.114.002450.
2) O'Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA, Granger CB, Krumholz HM, Linderbaum JA, Morrow DA, Newby LK, Ornato JP, Ou N, Radford MJ, Tamis-Holland JE, Tommaso CL, Tracy CM, Woo YJ, Zhao DX. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013;61(4):e78-140.
3) Rathore SS, Curtis JP, Chen J, et. al. Association of door-to-balloon time and mortality in patients admitted to hospital with ST elevation myocardial infarction: national cohort study. BMJ : British Medical Journal. 2009;338:b1807. doi:10.1136/bmj.b1807.

1a.3.1. What is the source of the systematic review of the body of evidence that supports the performance measure?
☒ Clinical Practice Guideline recommendation – ***complete sections 1a.4, and 1a.7***
☐ US Preventive Services Task Force Recommendation – ***complete sections 1a.5 and 1a.7***
☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*) – ***complete sections 1a.6 and 1a.7***
☒ Other – ***complete section 1a.8***

Please complete the sections indicated above for the source of evidence. You may skip the sections that do not apply.

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION
1a.4.1. Guideline citation (*including date*) and **URL for guideline** (*if available online*):

Antman EM, Anbe DT, Armstrong PW, Bates ER, Green LA, Hand M, Hochman JS, Krumholz HM, Kushner FG, Lamas GA, Mullany CJ, Ornato JP, Pearle DL, Sloan MA, Smith SC Jr. ACC/AHA guidelines for the management of patients with ST-elevation myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1999 Guidelines for the Management of Patients With Acute Myocardial Infarction). 2004.

The clinical practice guideline provided is based on its relevance to the measure. The guideline, released in 2013 by the American College of Cardiology Foundation (ACCF) and the American Heart Association (AHA), evaluates management of patients with STEMI. Citation for the guideline follows:

O'Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA, Granger CB, Krumholz HM, Linderbaum JA, Morrow DA, Newby LK, Ornato JP, Ou N, Radford MJ, Tamis-Holland JE, Tommaso CL, Tracy CM, Woo YJ, Zhao DX. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013 Jan 29;61(4):e78-140. Guideline available at: http://content.onlinejacc.org/article.aspx?articleid=1486115

1a.4.2. Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

"if PCI is chosen, the delay from patient contact with the healthcare system (typically, arrival at the ED or contact with paramedics) to balloon inflation should be less than 90 minutes." Page 593

The American College of Cardiology Foundation (ACCF)/American Heart Association (AHA) guideline for the management of ST-elevation myocardial infarction provides recommendations for the transfer of patients who require primary PCI

(pPCI), from a non-PCI-capable hospital to a PCI-capable hospital, in cases where pPCI can be performed within 120 minutes of first medical contact (FMC). Two recommendations support the measure's clinical intent:

A. Immediate transfer to a PCI-capable hospital for primary PCI is the recommended triage strategy for patients with STEMI who initially arrive at or are transported to a non-PCI capable hospital, with an FMC-to-device time system goal of 120 minutes or less. (Class I, Level of Evidence: B; pg. e86)
B. Immediate transfer to a PCI-capable hospital for coronary angiography is recommended for suitable patients with STEMI who develop cardiogenic shock or acute severe HF, irrespective of the time delay from MI onset. (Class I, Evidence: B; pg. e97)

1a.4.3. Grade assigned to the quoted recommendation with definition of the grade:

B ABC Scale

The recommendations received a Class I, Level B designation, indicating that the transfer of STEMI patients who require PCI to a PCI-capable hospital should occur immediately when the anticipated time from FMC-to-balloon is expected to be less than 120 minutes, and that patients who develop cardiogenic shock or acute severe heart failure (HF) should immediately be transferred to a PCI-capable hospital. These are strong recommendations supported by the evidence. The grade level of *B* does not indicate that the evidence supporting the guideline is weak, as the clinical questions addressed in the guideline do not lend itself to clinical trials.

The following grading scale applies to recommendations from the guideline:
Recommendation A: Class I: Benefit >>> Risk – Procedure/Treatment should be performed/administered
Recommendation B: Class I: Benefit >>> Risk – Procedure/Treatment should be performed/administered

The following evidence scale applies to recommendations from the guideline:
Recommendation A: Level B: Data derived from a single randomized trial or nonrandomized studies
Recommendation B: Level B: Data derived from a single randomized trial or nonrandomized studies

1a.4.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: If separate grades for the strength of the evidence, report them in section 1a.7.*)

- Level A (randomized controlled trial/ meta-analysis):

High quality randomized controlled trial that considers all important outcomes. High-quality meta-analysis (quantitative systematic review) using comprehensive search strategies.

- Level B (other evidence):

A well-designed, nonrandomized clinical trial. A nonquantitative systematic review with appropriate search strategies and well-substantiated conclusions. Includes lower quality randomized controlled trials, clinical cohort studies, and case-controlled studies with nonbiased selection of study participants and consistent findings. Other evidence, such as high-quality, historical, uncontrolled studies, or well-designed epidemiologic studies with compelling findings, is also included.

- Level C (consensus/expert opinion):

Consensus viewpoint or expert opinion. Expert opinion is sometimes the best evidence available.

Additional grading scale for the recommendations:
Class IIa: Benefit >>Risk additional studies with focused objectives needed. ***It is reasonable*** *to perform/administer treatment*
Class IIb: Benefit ≥ Risk additional studies with broad objectives needed: additional registry data would be helpful. Procedure/Treatment ***may be considered****.*
Class III No Benefit: Procedure/Test: not helpful, Treatment: no proven benefit
Class III Harm: Procedure/Test: excess cost w/o benefit or harmful, Treatment: harmful to patients

Additional evidence scales:
Level A: Multiple populations evaluated. Data derived from multiple randomized clinical trials or meta-analyses.
Level C: Very limited populations evaluated. Only consensus opinions of experts, case studies, or standard of care.

1a.4.5. Citation and URL for methodology for grading recommendations (*if different from 1a.4.1*)**:**

ACCF/AHA Task Force on Practice Guidelines. Manual for ACCF/AHA Guideline Writing Committees: Methodologies and Policies from the ACCF/AHA Task Force on Practice Guidelines. American College of Cardiology and American Heart Association. 2006. Available at: http://assets.cardiosource.com/Methodology_Manual_for_ACC_AHA_Writing_Committees.pdf and http://my.americanheart.org/professional/StatementsGuidelines/PoliciesDevelopment/Development/Methodologies-and-Policies-from-the-ACCAHA-Task-Force-on-Practice-Guidelines_UCM_320470_Article.jsp.

1a.4.6. If guideline is evidence-based (rather than expert opinion), are the details of the quantity, quality, and consistency of the body of evidence available (e.g., evidence tables)?

☒ Yes → ***complete section 1a.7***

☐ No → ***report on another systematic review of the evidence in sections 1a.6 and 1a.7; if another review does not exist, provide what is known from the guideline review of evidence in 1a.7***

1a.5. UNITED STATES PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

1a.5.1. Recommendation citation (*including date*) and **URL for recommendation** (*if available online*):

This measure is not based on a United States Preventive Services Task Force Recommendation.

1a.5.2. Identify recommendation number and/or page number and **quote verbatim, the specific recommendation**.

This measure is not based on a United States Preventive Services Task Force Recommendation.

1a.5.3. Grade assigned to the quoted recommendation with definition of the grade:

This measure is not based on a United States Preventive Services Task Force Recommendation.

1a.5.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: the grading system for the evidence should be reported in section 1a.7.*)

This measure is not based on a United States Preventive Services Task Force Recommendation.

1a.5.5. Citation and URL for methodology for grading recommendations (*if different from 1a.5.1*)**:**

This measure is not based on a United States Preventive Services Task Force Recommendation.

Complete section 1a.7

1a.6. OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1. Citation (*including date*) and **URL** (*if available online*):

Guidelines are evidenced based; details are provided in **section 1a.7**.

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

Guidelines are evidenced based; details are provided in **section 1a.7**.

Complete section 1a.7

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

If more than one systematic review of the evidence is identified above, you may choose to summarize the one (or more) for which the best information is available to provide a summary of the quantity, quality, and consistency of the body of evidence. Be sure to identify which review is the basis of the responses in this section and if more than one, provide a separate response for each review.

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

Methodologic Approach for the Systematic Review that Supports the Guideline

Members of the writing committee were appointed by the ACCF/AHA Task Force on Practice Guidelines (Task Force) and selected from the American College of Physicians, American College of Emergency Physicians, and Society for Cardiovascular Angiography and Interventions, representing various areas of medical expertise. Strict adherence to the Task Force Relationship with Industry (RWI) policy was maintained throughout the consensus process. The focus of the guideline is the management of patients with STEMI; particular emphasis has been placed in areas such as reperfusion therapy and organization of regional systems of care. Panel members extensively reviewed the relevant literature focusing on publications through November 2010, with additional selected references added through August 2012. Evidence supporting each guideline recommendation was weighted and ranked against the ACCF/AHA grading system. Recommendations have been developed using evidence-based methodologies created by the Task Force.

1a.7.2. Grade assigned for the quality of the quoted evidence with definition of the grade:

Grade for the evidence provided from the guideline can be found in **section 1a.4.3.**

1a.7.3. Provide all other grades and associated definitions for strength of the evidence in the grading system.

Grade for the evidence provided from the guideline can be found in **section 1a.4.3.**

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range:

It is inferred that the time period covered by the body of evidence is 2002-2012, as this is the period covered by the evidence cited for recommendations A and B.

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

The guideline does not explicitly indicate the specific number or type of study designs included in the body of evidence; however, the recommendations are Level B, and are based on data from a single randomized trial or non-randomized studies. Recommendation A references four unique citations, with evidence from one randomized trial, and one meta-analysis. Recommendation B references one unique citation with evidence from a single randomized trial.

1a.7.6. What is the overall quality of evidence across studies in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

The recommendations from the guideline are Class I, indicating that the benefits clearly outweigh the risks and the recommendations can be applied to most patients in most circumstances. The two Level B recommendations are based

on randomized control trials and a meta-analysis. The evidence presented has no important limitations and further evidence is unlikely to change the confidence in the estimate of the the effect.

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) across studies in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

The guideline does not provide details about the estimates of benefit and consistency across studies in the body of evidence.

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

The guideline does not provide details about potential harms associated with the timely transfer of STEMI patients requiring a PCI.

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for each new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

In addition to the guideline cited above, a review of the clinical literature was conducted during the measure contractor's annual review of the literature for additional evidence and/or new studies that relate to the measure. Citations and summaries for five items included in this review can be found in **section 1a.8.2**. Some of these five studies have been published since the period of guideline development. Results cited in these studies are consistent across studies and with the guidelines cited above.

1a.8 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.8.1 What process was used to identify the evidence?

In addition to the guideline cited above, a review of the clinical literature and related policy was conducted during the measure contractor's annual review of the literature for additional evidence and/or new studies that support the measure's intent. The measure contractor identified relevant peer-reviewed publications by searching the PubMed MEDLINE database from January 1, 2013 to September 1, 2015, limiting included results to those published in the English language and that had abstracts available in PubMed. The search initially identified 103 articles; a further review by the contractor's clinical and measure-development team resulted in the inclusion of five articles in the body of evidence below. Citations and summaries for the five items included in this review can be found in **section 1a.8.2**.

1a.8.2. Provide the citation and summary for each piece of evidence.

Anderson LL, French WJ, Peng SA, Vora AN, Henry TD, Roe MT, Kontos MC, Granger CB, Bates ER, Hellkamp A, and Wang TY. Direct transfer from the referring hospitals to the catheterization laboratory to minimize reperfusion delays for primary percutaneous coronary intervention insights from the national cardiovascular data registry. Circulation: Cardiovascular Interventions. 2015; 8(9): e002477. doi: 10.1161/CIRCINTERVENTIONS.114.002477.

Anderson et al. studied 33,901 patients with STEMI transferred for pPCI in the Acute Coronary Treatment and Intervention Outcomes Network Registry—Get With The Guidelines from July 2008 to December 2012. The majority of patients (78.2%) were transferred directly to the catheterization laboratory (cath lab), while the remaining were transferred first to the emergency department. The study found patients transferred directly to the cath lab had significantly lower door-to-balloon time compared to the patients transferred first to the

emergency department. A multivariable logistic regression further revealed transferring directly to the cath lab was associated with significantly faster reperfusion and lower mortality risk (odds ratio 0.58, 95% confidence interval 0.51-0.66, P<0.0001).

Dauerman HL, Bates ER, Kontos MC, Li S, Garvey JL, Henry TD, Manoukian SV, Roe MT. Nationwide analysis of patients with ST-segment–elevation myocardial infarction transferred for primary percutaneous intervention: Findings from the American Heart Association mission: Lifeline program. Circulation: Cardiovascular Interventions. 2015; 8(5): e002450. doi: 10.1161/ CIRCINTERVENTIONS.114.002450.

Dauerman et al. identified hospital-level, patient-level, and process characteristics of timely versus delayed PCI from a diverse national sample. Timely initiation of PCI is based on the American College of Cardiology/American Heart Association recommendation of door-to-device times within 120 minutes for transfer patients. Patients with a transfer time greater than 60 minutes were excluded from the study. 65% of patients met the recommended door-to-device time </= 120 minutes. Only 37% of the hospitals were high-performing (the study identifies high performing hospitals as having 75% or more of transfer STEMI patients with a first door-to-balloon time within 120 minutes). In addition to known predictors of delay, delays observed in this study were attributed to STEMI referral hospitals' rural location, longer estimated transfer time, and lower annual PCI hospital STEMI volumes.

Simon EL, Griffin P, Medepalli K, Griffin G, Williams CJ, Hewit M, Lloyd TS. Door-to-balloon times from freestanding emergency departments meet ST-segment elevation myocardial infarction reperfusion guidelines. The Journal of Emergency Medicine. 2015; 46(5): 734-740. doi: 10.1016/ j.jemermed.2013.08.089.

Simon et al. conducted a dual-center retrospective cohort review of all patients 18 years and older who were diagnosed with STEMI and presented to the main hospital-affiliated freestanding emergency departments (FEDs) from July 2007 to August 2008. The purpose of this study was to determine the proportion of STEMI patients who arrived to a FED and were subsequently transferred for PCI and met the door-to-balloon reperfusion guidelines of 90 minutes. There were 47 patients that met the study inclusion criteria. The median door-to-balloon time was 83 minutes. Of the 47 patients included in the study, 78.7% of the patients achieved a door-to-balloon time of 90 minutes or less.

Thilo C, Blüthgen A, and von Scheidt W. Efficacy and limitations of a STEMI network: 3 years of experience within the myocardial infarction network of the region of Augsburg-HERA. Clinical Research in Cardiology. 2013; 102(12): 905-914. doi: 10.1007/s00392-013-0608-8.

Thilo et al used the HERA Registry to investigate logistics, adherence to standards, time intervals, and mortality for pPCI in STEMI patients in a mixed urban and rural area. The study was comprised of 826 consecutive patients and 143 (17.3 %) patients from the sample presented in cardiogenic shock. Six hundred and eighty patients (82 %) received pPCI and 45 patients (5 %) received acute bypass surgery. For patients receiving pPCI, in-hospital mortality was 6.2 %, 28.0 % for shock patients, and 2.1 % for non-shock patients. The median first medical contact-to-balloon time was 135 minutes and median door-to-balloon time was 70 minutes. For patients whose first medical contact was by an emergency physician, when stratified by location, the median door-to-balloon time was 38 minutes with direct transfer to cath lab (n = 70), 69 minutes with direct transfer to the ICU (n = 240), and 132 minutes with direct transfer to the ER (n = 91, p < 0.01). The study concluded direct transfer to cath lab reduces door-to-balloon times by 49 %.

Vora AN, Holmes DN, Rokos I, Roe MT, Granger CB, French WJ, Antman E, Henry TD, Thomas L, Bates ER and Wang TY. Fibrinolysis use among patients requiring interhospital transfer for ST-segment elevation myocardial infarction care a report from the US national cardiovascular data registry. JAMA Internal Medicine. 2015; 175(2): 207-215. doi: 10.1001/jamainternmed.2014.6573.

Vora et al. assessed the association of estimated interhospital drive times with reperfusion strategy selection among the 22,481 transferred patients with STEMI in the United States between July 1, 2008, and March 31, 2012. 42.6% of transfer patients treated with pPCI satisfied the first door-to-balloon time of 120 minutes when the interhospital driving time surpassed 30 minutes. 52.7% of eligible patients with drive times exceeding 60 minutes received fibrinolysis. Of the patients eligible for fibrinolysis or pPCI with driving times ranging from 30-

120 minutes, 34.3% received pretransfer fibrinolysis, with a median door-to-needle time of 34 minutes. Although there was not a significant mortality difference for patients treated with pPCI compared to those treated with fibrinolysis (3.7% vs 3.9%; adjusted odds ratio, 1.13; 95% CI, 0.94-1.36), there was a higher bleeding risk for patients who received fibrinolysis (10.7% vs 9.5%; adjusted odds ratio, 1.17; 95% CI, 1.02-1.33). The authors concluded neither fibrinolysis nor pPCI were being adequately administered within guideline-recommended reperfusion targets. In addition, the authors concluded that patients who are unlikely to receive timely pPCI should receive pretransfer fibrinolysis, followed by early transfer for angiography as an alternative in situations where potential benefits of timely reperfusion outweigh bleeding risk.

1. Evidence, Performance Gap, Priority – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all subcriteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form
NQF_0290_MeasureEvidenceForm.docx

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- disparities in care across population groups.

1b.1. **Briefly explain the rationale for this measure** *(e.g., the benefits or improvements in quality envisioned by use of this measure)*
The early use of primary angioplasty in patients with STEMI results in a significant reduction in mortality and morbidity. The earlier PCI is provided, the more effective it is (Rathore, 2009). National guidelines recommend the prompt initiation of PCI in patients presenting with STEMI infarction (O'Gara, 2013). Current recommendations support a door-to-balloon time of 120 minutes or less in patients that need to be transferred from a non-PCI capable hospital to a PCI-capable hospital (O'Gara, 2013). Patients transferred for primary PCI (pPCI) rarely meet recommended guidelines for door-to-balloon time (Dauerman et. al., 2015). Elevated transfer time is a significant predictor of delay in the initiation of PCI (Dauerman et. al., 2015). Decreasing transfer time in STEMI patients requiring an acute coronary intervention has the potential to lead to reduced door-to balloon-time.

REFERENCES:
1) Dauerman HL, Bates ER, Kontos MC, Li S, Garvey JL, Henry TD, Manoukian SV, Roe MT. Nationwide analysis of patients With ST-segment-elevation myocardial infarction transferred for primary percutaneous intervention: findings from the american heart association mission: lifeline program. Circ Cardiovasc Interv. 2015:8(5):pii: e002450. doi: 10.1161/CIRCINTERVENTIONS.114.002450.
2) O´Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA, Granger CB, Krumholz HM, Linderbaum JA, Morrow DA, Newby LK, Ornato JP, Ou N, Radford MJ, Tamis-Holland JE, Tommaso CL, Tracy CM, Woo YJ, Zhao DX. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2013;61(4):e78-140.
3) Rathore SS, Curtis JP, Chen J, et. al. Association of door-to-balloon time and mortality in patients admitted to hospital with ST elevation myocardial infarction: national cohort study. BMJ?: British Medical Journal. 2009;338:b1807. doi:10.1136/bmj.b1807.

1b.2. **Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.** *(This is required for endorsement maintenance. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included). This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*
Analysis of facility-level data from the Hospital Compare downloadable files indicates that there is variation in the median time to transfer for ACI for cases where patients have a principal diagnosis associated with AMI and have ST-segment elevation on the ECG closest to ED arrival. During the April 2010-March 2011 data collection period, median transfer times ranged from 26 minutes to 245 minutes with a median of 56 minutes. During the April 2014-March 2015 data collection period, median transfer times ranged from 20 minutes to 474 minutes, with a median of 54 minutes. From the March 2010-April 2015 data collection periods, the median value of median time to transfer for ACI for cases where patients have a principal diagnosis associated with AMI and have ST-segment

elevation on the ECG closest to ED arrival decreased 3.6% (-2 minutes).

The data presented below represent performance scores and descriptive statistics for longitudinal facility performance for the facilities whose denominator counts met minimum case count requirements during the April 2010-March 2015 data collection periods.

	Data Collection Period		Change in Minutes			
	April 2010-March 2011		April 2011-March 2012		April 2012-March 2013	April 2013-March 2014	April 2014-March 2015	2010-2014
Facilities	421	400	405	409	425	-		
Minimum Value	26	20	14	21	20	-6		
1st Percentile	26	22	24	25	26	-		
5th Percentile	32	31	30	30	30	-2		
10th Percentile	36	35	34	34	33	-3		
25th Percentile	45	44	42	41	42	-3		
Median	56	54	54	54	54	-2		
75th Percentile	70	68	68	71	69	-1		
90th Percentile	91	90	89	87	94	3		
95th Percentile	104	120	110	108	113	9		
99th Percentile	210	229	152	166	206	-4		
Maximum Value	245	542	307	288	474	229		
Number of cases who were transferred for ACI (Denominator)	8,008	7,621	7,822	7,678	8,166	-		

From the inception of public reporting through March 2015 data collection, there has been wide variation in facility performance. The interquartile range has been consistently wide, ranging from 24 minutes to 30 minutes. Additionally, the maximum time for transfer for ACI varies greatly over the five years of data collection. While median performance is improving, there is an ongoing opportunity for improvement in facility performance.

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.
Data have been included in Section 1b.2; these data represent national performance over time, from April 2010-March 2015 data collection periods.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for endorsement maintenance. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*
The relationship of patient and facility characteristics on time to transfer from the ED to another facility for ACI was evaluated using 2014 data submitted to the Clinical Data Warehouse (CDW).

Primary results from the regression were related to patient demographics. There was a significantly lower time to transfer for ACI for patients in age groups 40-50 (ß= -17.5 minutes, p= 0.02) and 50-60 (ß= -16.9 minutes, p= 0.03), as compared to patients in age group 18-30. There was a significantly longer time to transfer for ACI for female patients, as compared to male patients (ß= 14.4 minutes, p< 0.001). In comparison to White patients, there was a significantly longer time to transfer for ACI for African American (ß= 25.0 minutes, p< 0.001) and Asian (ß= 11.5 minutes, p= 0.04) patients. Additionally, there was a significantly longer time to transfer for ACI for Hispanic patients, as compared to non-Hispanic patients (ß= 23.6 minutes, p= 0.01).

Time to transfer for ACI also varied by the characteristics of the facility that the patient was transferred from. When compared to patients transferred from facilities with fewer than 50 beds (a proxy for facility size), there was a significantly lower time to transfer for ACI for patients transferred from facilities with 101-250 beds (ß= -13.6 minutes, p <0.001), and 500+ beds (ß= -20.3 minutes, p< 0.001). There was a significantly lower time to transfer for ACI for patients transferred from an urban hospital as compared to patients transferred from a rural hospital (ß= -15.7 minutes, p< 0.001). Finally, when compared to patients transferred from a non-teaching facility, there was a significantly longer time to transfer for ACI for patients transferred from a teaching facility (ß= 8.7 minutes, p= 0.02) and a major teaching facility (ß= 37.4 minutes, p< 0.001).

1b.5. **If no or limited data on disparities from the measure as specified is reported in 1b4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations.**
Results of the linear regression (Section 1b.4) identified statistically significant differences in time to transfer for ACI for several patient and facility level characteristics, including sex. This difference is also supported in the literature. Evidence in the literature demonstrates that women who require PCI are more likely to exceed transfer guidelines for PCI when compared to men (odds ratio: 2.63; 95% confidence interval: 1.17 – 4.07). This relationship between sex and transfer time persists after adjusting for other sociodemographic, clinical, and organizational factors. (D'Onofrio et al., 2015)

REFERENCES:
1) D'Onofrio G, Safdar B, Lichtman JH, et al. Sex Differences in Reperfusion in Young Patients With ST-Segment–Elevation Myocardial Infarction Results From the VIRGO Study. Circulation. 2015;131(15):1324-1332. doi:10.1161/CIRCULATIONAHA.114.012293.

1c. High Priority (previously referred to as High Impact)

The measure addresses:

- a specific national health goal/priority identified by DHHS or the National Priorities Partnership convened by NQF; OR
- a demonstrated high-priority (high-impact) aspect of healthcare (e.g., affects large numbers of patients and/or has a substantial impact for a smaller population; leading cause of morbidity/mortality; high resource use (current and/or future); severity of illness; and severity of patient/societal consequences of poor quality).

1c.1. **Demonstrated high priority aspect of healthcare**
Affects large numbers, A leading cause of morbidity/mortality, Patient/societal consequences of poor quality
1c.2. **If Other:**

1c.3. **Provide epidemiologic or resource use data that demonstrates the measure addresses a high priority aspect of healthcare. List citations in 1c.4.**
According to the Centers for Disease Control and Prevention (CDC), heart disease is the leading cause of death in the United States (CDC, 2015). Approximately 735,000 Americans develop AMI each year, of which 525,000 are incident cases (Mozaffarian et. al., 2015). National guidelines suggest that timely initiation of PCI is the recommended treatment strategy in patients that present with myocardial infarction (O'Gara 2013), however, only 39% of US hospitals are PCI-capable (Langabeer, 2013), so transferring patients is the most viable option if the suspected door-to-balloon time is within 120 minutes (O'Gara, 2013). Evidence shows that patients who receive PCI within the recommended guidelines have a reduced risk of morbidity and mortality (Rathore, 2009). Despite clear evidence that timely initiation of PCI is associated with improved health outcomes, a study by Dauerman and colleagues (2015) demonstrated that more than one-third of patients transferred for PCI fail to meet the recommended door-to-balloon time of 120 minutes or less. While the time to PCI at the receiving facility is beyond the control of facilities reporting the measure, timely transfer can help to reduce the overall door-to-balloon time, resulting in improved health outcomes.

To complement a review of the literature, national statistics on ED use from the Healthcare Cost and Utilization Project (HCUP) National Emergency Department Sample (NEDS) were extracted to estimate the total patient population of patients with a principal diagnosis associated with AMI. In 2013, there were an estimated 546,352 ED visits with a primary diagnosis of AMI. Most patients were ages 65-84 (42.3%), followed by patients ages 45-64 (37.2%), 85+ (14.2%), 18-44 (5.6%), and finally patients under 18 (0.7%). There were more male patients (61.1%) than female patients (39.9%).

Additionally, the Department of Health and Human Services (DHHS) includes AMI care in a number of national programs, including the Million Hearts Campaign and Healthy People 2020. Both of these initiatives galvanize existing efforts and new programs to improve cardiovascular health and quality of life through prevention, detection, and treatment of AMI and strokes (Frieden and Berwick 2011; DHHS 2014). Reducing door-to-balloon time through the continued reporting of NQF #0290 can help DHHS achieve its Healthy People 2020 objective of increasing the proportion of eligible patients with AMI who receive timely artery-opening therapy as specified by current guidelines (DHHS, 2014).

The literature and utilization data demonstrate that AMI treatment remains a high priority aspect of healthcare because AMI is a leading cause of morbidity and mortality that affects a large number of individuals, with severe consequences resulting from poor quality of care.

1c.4. **Citations for data demonstrating high priority provided in 1a.3**
1) CDC, NCHS. Underlying Cause of Death 1999-2013 on CDC WONDER Online Database, released 2015. Data are from the

Multiple Cause of Death Files, 1999-2013, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Accessed Feb. 3, 2015.
2) Dauerman HL, Bates ER, Kontos MC, Li S, Garvey JL, Henry TD, Manoukian SV, Roe MT. Nationwide analysis of patients With ST-segment-elevation myocardial infarction transferred for primary percutaneous intervention: findings from the american heart association mission: lifeline program. Circ Cardiovasc Interv. 2015:8(5):pii: e002450. doi: 10.1161/CIRCINTERVENTIONS.114.002450.
3) DHHS. 2014. HDS-19.3 (Developmental) Increase the proportion of eligible patients with strokes who receive acute reperfusion therapy within 3 hours from symptom onset. http://www.healthypeople.gov/node/4581/data_details.
4) Frieden TR and Berwick D. The "Million Hearts" initiative – preventing heart attacks and strokes. N Engl J Med. 2011;365:e27.
5) Langabeer JR, Henry TD, Kereiakes DJ, Dellifraine J, Emert J, Wang Z, Stuart L, King R, Segrest W, Moyer P, Jollis JG. Growth in percutaneous coronary intervention capacity relative to population and disease prevalence. J Am Heart Assoc. 2013;2(6):e000370. doi: 10.1161/JAHA.113.000370.
6) Mozaffarian D, Benjamin EJ, Go AS, et. al. Heart disease and stroke statistics—2015 update: a report from the American Heart Association. Circulation. 2015;131:e29-322.
7) O´Gara PT, Kushner FG, Ascheim DD, Casey DE Jr, Chung MK, de Lemos JA, Ettinger SM, Fang JC, Fesmire FM, Franklin BA, Granger CB, Krumholz HM, Linderbaum JA, Morrow DA, Newby LK, Ornato JP, Ou N, Radford MJ, Tamis-Holland JE, Tommaso CL, Tracy CM, Woo YJ, Zhao DX. 2013 ACCF/AHA guideline for the management of ST-elevation myocardial infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J 8)
Am Coll Cardiol. 2013;61(4):e78-140.
8) Rathore SS, Curtis JP, Chen J, et. al. Association of door-to-balloon time and mortality in patients admitted to hospital with ST elevation myocardial infarction: national cohort study. BMJ?: British Medical Journal. 2009;338:b1807. doi:10.1136/bmj.b1807

1c.5. **If a PRO-PM (e.g. HRQoL/functional status, symptom/burden, experience with care, health-related behaviors), provide evidence that the target population values the measured PRO and finds it meaningful.** *(Describe how and from whom their input was obtained.)*
This measure is not a PRO-PM measure.

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the subcriteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. **Subject/Topic Area** *(check all the areas that apply):*
Cardiovascular, Cardiovascular : Acute Myocardial Infarction

De.6. **Cross Cutting Areas** *(check all the areas that apply):*
Care Coordination, Safety, Safety : Complications

S.1. **Measure-specific Web Page** *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*
https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier2&cid=1196289981244

S.2a. **If this is an eMeasure**, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)
This is not an eMeasure **Attachment:**

S.2b. **Data Dictionary, Code Table, or Value Sets** *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*
Attachment **Attachment:** NQF_0290_MeasureCodeSet.xlsx

S.3. **For endorsement maintenance**, please briefly describe any changes to the measure specifications since last endorsement date and explain the reasons.
NQF #0290 was first endorsed by NQF in November 2007. Since 2007, the measure specifications have been updated to reflect clinical changes in appropriate AMI diagnosis or ED evaluation and management (E/M) codes; to address stakeholder feedback; and to harmonize with measures in the Hospital Inpatient Quality Reporting (HIQR) program. Some data elements have been updated to provide clarification in abstraction and updates have been made to selected references since the measure's most recent NQF review, in 2012. The Discharge Status data element was changed to Discharge Code. Multiple clarifications and modifications were made to the Initial ECG Interpretation data element to eliminate false inclusions and exclusions, and the list of negative qualifiers and modifiers was updated to align with a measure reported in the HIQR Program. The Reason for Not Administering Fibrinolytic Therapy data element was modified to include a pre-arrival list of medication that would count as an automatic reason for not administering fibrinolytic therapy.

In 2014, left bundle branch block (LBBB) was removed from the measure as an inclusionary ECG finding in order to align with updates made in the 2013 STEMI guidelines (O'Gara, 2013).

In 2015, as part of the annual measure maintenance and review process, all ICD-9-CM diagnosis codes were updated to corresponding ICD-10-CM diagnosis codes.

S.4. **Numerator Statement** *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome)*
IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
This measure is reported as a continuous variable statement: time (in minutes) from ED arrival to transfer to another facility for ACI.

The numerator includes patients with AMI and ST-segment elevation on the ECG performed closest to ED arrival who are transferred from the ED to a short-term general hospital for inpatient care, or to a Federal healthcare facility specifically for ACI.

S.5. **Time Period for Data** *(What is the time period in which data will be aggregated for the measure, e.g., 12 mo, 3 years, look back to August for flu vaccination? Note if there are different time periods for the numerator and denominator.)*
Median times to transfer within a three-month period are aggregated, on a rolling basis, for AMI patients who are transferred for ACI.

S.6. **Numerator Details** *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
The measure population is defined by six E/M codes and 18 ICD-10-CM diagnosis codes included in the code set for this measure; these detailed lists can be found in the Excel workbook provided for Section S.2b.

The measure population includes patients with AMI and ST-segment elevation on the ECG performed closest to ED arrival who are transferred from the ED to a short-term general hospital for inpatient care, or to a Federal healthcare facility specifically for an acute coronary intervention. Patients are included in the measure population if:

- Initial ECG Interpretation is equal to "Yes";
- Fibrinolytic Administration is equal to "No"; and
- Transfer for Acute Coronary Intervention is equal to "[1] There was documentation the patient was transferred from this facility's emergency department to another facility specifically for acute coronary intervention."

S.7. **Denominator Statement** *(Brief, narrative description of the target population being measured)*
Time (in minutes) from ED arrival to transfer to another facility for ACI.

S.8. **Target Population Category** *(Check all the populations for which the measure is specified and tested if any):*
Populations at Risk, Senior Care

S.9. **Denominator Details** *(All information required to identify and calculate the target population/denominator such as definitions, specific data collection items/responses , code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should*

be provided in an Excel or csv file in required format at S.2b)
NQF #0290 is a continuous measure; therefore, the numerator and denominator details contained in Section S.6 and Section S.9 are the same.

The measure population is defined by six E/M codes and 18 ICD-10-CM diagnosis codes included in the code set for this measure; these detailed lists can be found in the Excel workbook provided for Section S.2b.

The measure population includes patients with AMI and ST-segment elevation on the ECG performed closest to ED arrival who are transferred from the ED to a short-term general hospital for inpatient care, or to a Federal healthcare facility specifically for ACI. Patients are included in the measure population if:

- Initial ECG Interpretation is equal to "Yes";
- Fibrinolytic Administration is equal to "No"; and
- Transfer for Acute Coronary Intervention is equal to "[1] There was documentation the patient was transferred from this facility's emergency department to another facility specifically for acute coronary intervention."

S.10. **Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*
Patients are excluded from this measure if they are under 18 years of age, did not have an initial ECG interpretation, received fibrinolytic therapy while in the ED, or were transferred for reasons other than ACI.

S.11. **Denominator Exclusion Details** *(All information required to identify and calculate exclusions from the denominator such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
Cases are excluded for any patients that meet any of the following criteria: Patients less than 18 years of age

- Initial ECG Interpretation is equal to "No"
- Fibrinolytic Administration is equal to "Yes"
- Transfer for Acute Coronary Intervention is equal to "[2] There was documentation the patient was admitted to observation." or "[3] There was documentation the patient was transferred from this facility's emergency department to another facility for reasons other than acute coronary intervention, or the specific reason for transfer was unable to be determined from medical record documentation."

S.12. **Stratification Details/Variables** *(All information required to stratify the measure results including the stratification variables, definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b)*
Not applicable; this measure does not stratify its results.

S.13. **Risk Adjustment Type** (Select type. Provide specifications for risk stratification in S.12 and for statistical model in S.14-15)
No risk adjustment or risk stratification
If other:

S.14. **Identify the statistical risk model method and variables** *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development and testing should be addressed with measure testing under Scientific Acceptability)*
Not applicable; this measure does not risk adjust.

S.15. **Detailed risk model specifications** *(must be in attached data dictionary/code list Excel or csv file. Also indicate if available at measure-specific URL identified in S.1.)*
Note: Risk model details (including coefficients, equations, codes with descriptors, definitions), should be provided on a separate worksheet in the suggested format in the Excel or csv file with data dictionary/code lists at S.2b.
Provided in response box S.15a

S.15a. **Detailed risk model specifications** (*if not provided in excel or csv file at S.2b*)
No risk model specifications are provided, as risk adjustment or stratification is not necessary for this measure.

S.16. **Type of score:**
Continuous variable
If other:

S.17. **Interpretation of Score** *(Classifies interpretation of score according to whether better quality is associated with a higher score,*

a lower score, a score falling within a defined interval, or a passing score)
Better quality = Lower score

S.18. **Calculation Algorithm/Measure Logic** *(Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.)*
This measure calculates the time (in minutes) from ED arrival to transfer to another facility for ACI. The patient population is determined from two algorithms; the AMI Hospital Outpatient Population algorithm as well as the OP-3 measure-specific algorithm:

1. Check E/M Code; if on Table 1.0 (in the Excel workbook provided for Section S.2b), proceed.
2. Check Discharge Code; include patients with discharge code of 4a or 4d.
3. Calculate Patient Age (Outpatient Encounter Date - Birthdate).
4. Check Patient Age; if >= 18, proceed.
5. Check ICD-10-CM Principal Diagnosis Code; if on Table 1.1 (in the Excel workbook provided for Section S.2b), proceed to the measure-specific algorithm.
6. Check Initial ECG Interpretation. If Initial ECG Interpretation equals YES, the case will proceed to Fibrinolytic Administration.
7. Check Fibrinolytic Administration. If Fibrinolytic Administration equals NO, the case will proceed to Transfer for Acute Coronary Intervention.
8. Check Transfer for Acute Coronary Intervention. If Transfer for Acute Coronary Intervention equals 1 (i.e., there is documentation the patient was transferred from this facility's emergency department to another facility specifically for ACI), the case will proceed to ED Departure Date.
9. Check ED Departure Date. If ED Departure Date equals Non-UTD Value, the case will proceed to ED Departure Time.
10. Check ED Departure Time. If ED Departure Time equals Non-UTD Value, the case will proceed to Arrival Time.
11. Check Arrival Time. If Arrival Time equals Non-UTD Value, the case will proceed to the Measurement Value.
12. Calculate the Measurement Value. Time in minutes is equal to the ED Departure Date and ED Departure Time (in minutes) minus the Outpatient Encounter Date and Arrival Time (in minutes).
13. Check the Measurement Value. If Measurement Value is greater than or equal to 0 minutes, the case will proceed to Reason for Not Administering Fibrinolytic Therapy.
14. Check Reason for Not Administering Fibrinolytic Therapy. If Reason for Not Administering Fibrinolytic Therapy equals 1, 2, or 3, the case will proceed to a Measure Category Assignment of D1, the OP-3a Overall Rate. Initialize the Measure Category Assignment for OP-3b and OP-3c equal to B. Do not change the Measure Category Assignment that was already calculated for the overall rate of OP-3a. Proceed to Reason for Not Administering Fibrinolytic Therapy.
15. Check Reason for Not Administering Fibrinolytic Therapy. If Reason for Not Administering Fibrinolytic Therapy equals 1 or 2, the case will proceed to a Measure Category Assignment of D2, the OP-3c Quality Improvement Rate. If Reason for Not Administering Fibrinolytic Therapy equals 3, the case will proceed to a Measure Category Assignment of D, the OP-3b Reporting Rate. Return to Transmission Data Processing Flow: Clinical in the Data Transmission Section.

S.19. **Calculation Algorithm/Measure Logic Diagram URL or Attachment** *(You also may provide a diagram of the Calculation Algorithm/Measure Logic described above at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
Available at measure-specific web page URL identified in S.1

S.20. **Sampling** *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*
IF a PRO-PM, identify whether (and how) proxy responses are allowed.
Sampling is a process of selecting a representative part of a population in order to estimate the hospital's performance without collecting data for its entire population. Using a statistically valid sample, a hospital can measure its performance in an effective and efficient manner. Sampling is a particularly useful technique for performance measures that require primary data collection from a source such as the medical record. Sampling should not be used unless the hospital has a large number of cases in the outpatient population because a fairly large number of cases are needed to achieve a representative sample of the population. For the purpose of sampling outpatient department quality measures, the terms "sample," "effective sample," and "case" are defined below:
• The "sample" is the fraction of the population that is selected for further study.
• "Effective sample" refers to the part of the sample that makes it into the denominator of an outpatient measure set. This is defined as the sample for an outpatient measure set minus all the exclusions and contraindications for the outpatient measure set in the sample.
• A "case" refers to a single record (or an encounter) within the population. For example, during the first quarter a hospital may have 100 patients who had a principal diagnosis associated with the OP-1, 2, 3, 4, and 5 measures. The hospital's outpatient population would include 100 cases or 100 outpatient records for these measures during the first quarter.

To obtain statistically valid sample data, the sample size should be carefully determined, and the sample cases should be randomly selected in such a way that the individual cases in the population have an equal chance of being selected. Only when the sample data truly represent the whole population can the sample-based performance outpatient measure set data be meaningful and useful. Each hospital is ultimately responsible for adhering to the sampling requirements outlined in this manual.

As a general rule/policy of CMS, providers are encouraged to submit as many cases as possible up to the entire population of cases if reasonably feasible. For example, if the raw data can be easily extracted from an existing electronic database or the abstraction burden is manageable, providers should consider submitting the entire population of cases that meet the initial selection criteria. Otherwise, a statistically valid sample can be selected.

S.21. **Survey/Patient-reported data** *(If measure is based on a survey, provide instructions for conducting the survey and guidance on minimum response rate.)*
IF a PRO-PM, specify calculation of response rates to be reported with performance measure results.
This measure does not use survey data.

S.22. **Missing data** (specify how missing data are handled, e.g., imputation, delete case.)
Required for Composites and PRO-PMs.
The measure does not make any adjustments for missing data. If data are missing, the case will proceed to Measure Category Assignment of X and will be rejected. While abstractors cannot submit missing data, they can submit a value of "UTD" for select data elements. Depending on the data element the case is then either excluded from the denominator or excepted from the numerator. Frequency and distribution of data with a value of "UTD" are reported in the attached Measure Testing Form.

S.23. **Data Source** *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*
If other, please describe in S.24.
Administrative claims, Electronic Clinical Data, Electronic Clinical Data : Electronic Health Record, Paper Medical Records

S.24. **Data Source or Collection Instrument** *(Identify the specific data source/data collection instrument e.g. name of database, clinical registry, collection instrument, etc.)*
IF a PRO-PM, identify the specific PROM(s); and standard methods, modes, and languages of administration.
An electronic data collection tool is made available from vendors or facilities can download the free CMS Abstraction & Reporting Tool (CART). Paper tools for manual abstraction, which are posted on www.QualityNet.org, are also available for the CART tool. These tools are posted on www.QualityNet.org.

S.25. **Data Source or Collection Instrument** *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
Available at measure-specific web page URL identified in S.1

S.26. **Level of Analysis** *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*
Facility, Population : National

S.27. **Care Setting** *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*
Hospital/Acute Care Facility
If other:

S.28. **COMPOSITE Performance Measure** - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*
Not applicable; this is not a composite measure.

2a. Reliability – See attached Measure Testing Submission Form
2b. Validity – See attached Measure Testing Submission Form
NQF_0290_MeasureTestingForm.docx

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b2-2b7)

Measure Number (*if previously endorsed*)**:** 0290

Measure Title: Median Time to Transfer to Another Facility for Acute Coronary Intervention
Date of Submission: 4/29/2016
Type of Measure:

☐ Composite – ***STOP – use composite testing form***	☐ Outcome (*including PRO-PM*)
☐ Cost/resource	☒ Process
☐ Efficiency	☐ Structure

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- **For all measures, sections 1, 2a2, 2b2, 2b3, and 2b5 must be completed.**
- **For outcome and resource use measures**, section **2b4** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b6** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b2-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 20 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address sociodemographic variables and testing in this form refer to the release notes for version 6.6 of the Measure Testing Attachment.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **PRO-PMs and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b2. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **PRO-PMs and composite performance measures**, validity should be demonstrated for the computed performance score.

2b3. Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion; [12]
AND
If patient preference (e.g., informed decision making) is a basis for exclusion
, there must be evidence that the exclusion impacts performance on the measure; in such cases, 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). [13]

2b4. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and sociodemographic factors) that influence the measured outcome and are present at start of care; [14,15] and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b5. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow

for **identification of statistically significant and practically/clinically meaningful [16] differences in performance**;
OR
there is evidence of overall less-than-optimal performance.

2b6. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b7. For **eMeasures, composites, and PRO-PMs** (or other measures susceptible to missing data), analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

Notes
10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements 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 of the measure score addresses precision of measurement (e.g., signal-to-noise).
11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality.
12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.
13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.
14. Risk factors that influence outcomes should not be specified as exclusions
15. 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 percent v. 75 percent) 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 less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE
Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.23*)	**Measure Tested with Data From:**
☒ abstracted from paper record	☒ abstracted from paper record
☒ administrative claims	☒ administrative claims
☐ clinical database/registry	☐ clinical database/registry
☒ abstracted from electronic health record	☒ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

a) Datasets used to define the sample:

- The initial patient population is identified using chart-abstracted data for a sample of emergency department (ED) encounters with at least one of the following Current Procedural Terminology (CPT) codes for evaluation and management (E/M): 99281, 99282, 99283, 99284, 99285, or 99291. The initial patient population includes cases for patients 18 years and older, as of the date of the encounter, with a principal diagnosis associated with an acute myocardial infarction, identified by using any of the following International Classification of Diseases version 9 (ICD-9) codes: 410.00, 410.01, 410.10, 410.11, 410.20, 410.21, 410.30, 410.31, 410.40, 410.41, 410.50, 410.51, 410.60, 410.61, 410.70, 410.71, 410.80, 410.81, 410.90, or 410.91.[1]

b) Datasets used to define the effective sample:

- The effective sample is identified using chart-abstracted data for a sample of cases for patients included in the initial patient population.

c) Datasets used to identify denominator exclusions:

- Denominator exclusions are identified using chart-abstracted data of cases for patients included in the effective sample. Denominator exclusions capture cases for patients where any of the following conditions are met:
    - *Initial ECG Interpretation* equal to "No"
    - *Fibrinolytic Administration* is equal to "Yes"
    - *Transfer for Acute Coronary Intervention* is equal to "[2] there was documentation the patient was admitted to observation status prior to transfer"
    - *Transfer to Acute Coronary Intervention* is equal to "[3] There was documentation the patient was transferred from this facility's emergency department to another facility for reasons other than acute coronary intervention, or the specific reason for transfer was unable to be determined from medical record documentation"
    - *Reason for Not Administering Fibrinolytic Therapy* is equal to "[1] Documented contraindication/reason"
    - *Reason for Not Administering Fibrinolytic Therapy* is equal to "[2] Cardiogenic Shock"
    - *Reason for Not Administering Fibrinolytic Therapy* is equal to "[2] There was documentation the patient was admitted to observation status prior to transfer"
    - *Reason for Not Administering Fibrinolytic Therapy* is equal to "[3] There was documentation the patient was transferred from this facility's emergency department to another facility for reasons other than acute coronary intervention, or the specific reason for transfer was unable to be determined from medical record documentation"

d) Datasets used to capture the numerator:

- The numerator is identified using chart-abstracted data of cases for patients included in the effective sample. The numerator is reported as a continuous variable statement and includes cases for patients where both of the following conditions are met:
    - *Initial ECG Interpretation* is equal to "Yes"
    - *Fibrinolytic Administration* is equal to "No"
    - *Transfer for Acute Coronary Intervention* is equal to "[1] There was documentation the patient was transferred from this facility's emergency department to another facility specifically for acute coronary intervention"

[1] As of October 1, 2015 the measure is calculated using ICD-10 codes. However, as these data will not be available in the CDW until 2016, testing was conducted on the most recent version of the measure using ICD-9 codes.

- *Reason for Not Administering Fibrinolytic Therapy* is equal to "[3] No documented contraindication/reason or Unable to determine (UTD)"

e) Datasets used to identify numerator exceptions:

- Numerator exceptions are identified using chart-abstracted data of cases for patients included in the effective sample. Numerator exceptions include cases of patients for whom any of the following conditions are met:
  - *ED Departure Date* is equal to "UTD"
  - *ED Departure Time* is equal to "UTD"
  - *Arrival Time* is equal to "UTD"

1.3. What are the dates of the data used in testing April 2010-March 2015

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (*must be consistent with levels entered in item S.26*)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☐ group/practice	☐ group/practice
☒ hospital/facility/agency	☒ hospital/facility/agency
☐ health plan	☐ health plan
☒ other: national	☒ other: national

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

The number of measured entities (hospital EDs) varies by testing type; see Section **1.7** for details.

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

The number of patients varies by testing type; see Section **1.7** for details.

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.

Reliability Testing
Currently undergoing validation through the CMS Clinical Data Abstraction Center

Reliability Testing:
Data Source: Denominator: Clinical Data Warehouse (CDW); Numerator: CDW; Exclusions: CDW; Exceptions: CDW [maintained by the Centers for Medicare & Medicaid Services (CMS)]
Dates: Denominator: January 1, 2014-December 31, 2014; Numerator: January 1, 2014-December 31, 2014; Exclusions: January 1, 2014-December 31, 2014; Exceptions: January 1, 2014-December 31, 2014;
Number of Facilities: 1,902[2]
Sample (denominator): 64,827
Effective Sample (denominator after exclusions): 13,195

[2] The CDW sample included 3,024 facilities; 1,902 facilities had cases for patients who were eligible for NQF #0290.

Level of Analysis: Case
Sampled Patient Characteristics: Gender (% Male): 68.5; Mean Age (Years): 62.3 (St. Dev.: 13.8); Race (% Minority): 9.7

Validity Testing
Currently undergoing validation through the CMS Clinical Data Abstraction Center

Validity Testing – *Data Element Validity*
Data Source: Denominator: CDW; Numerator: CDW; Exclusions: CDW; Exceptions: CDW [maintained by CMS]
Dates: Denominator: April 1, 2014-March 31, 2015; Numerator: April 1, 2014-March 31, 2015; Exclusions: April 1, 2014-March 31, 2015; Exceptions: April 1, 2014-March 31, 2015
Number of Facilities: 65[3]
Sample (denominator): 462
Effective sample (denominator after exclusions): 86
Level of Analysis: Data element
Sampled Patient Characteristics: Gender (% Male): 63.4; Mean Age (Years): 62.8 (St. Dev.: 14.6); Race (% Minority): 10.7

Validity Testing – *Face Validity*
Data Source: Structured qualitative survey completed by the stroke and acute myocardial infarction expert work group (EWG) members
Date Collected: February 2016
Number of Responses: 5
Respondent Characteristics: Respondents were asked to self-identify as one or more of the following categories: insurer/purchaser; payer; clinician (5); management (1); healthcare administration (1); patient or patient advocate; caregiver.

Exclusions Analysis
N/A

Exclusions Analysis
Data Source: Denominator: CDW; Numerator: CDW; Exclusions: CDW; Exceptions: CDW [maintained by CMS]
Dates: Denominator: January 1, 2014-December 31, 2014; Numerator: January 1, 2014-December 31, 2014; Exclusions: January 1, 2014-December 31, 2014; Exceptions: January 1, 2014-December 31, 2014;
Number of Facilities: 1,902
Sample (denominator): 64,827
Effective Sample (denominator after exclusions): 13,195
Level of Analysis: Case
Sampled Patient Characteristics: Gender (% Male): 68.5; Mean Age (Years): 62.3 (St. Dev.: 13.8); Race (% Minority): 9.7

Risk Adjustment Strategy
N/A

Risk Adjustment/Stratification
N/A- No risk adjustment or stratification was performed.

Identification of Meaningful Differences in Performance
N/A

Identification of Statistically Significant & Meaningful Differences in Performance
Data Source: Hospital Compare downloadable file [maintained by CMS]

[3] The cases used for data element validity testing were abstracted from the CDW by CDAC and represent a sample of all CDW cases. The CDAC abstraction sample included 196 facilities; 65 facilities had cases for patients who were eligible for NQF #0290.

Dates: Denominator: April 1, 2014-March 31, 2015; Numerator: April 1, 2014-March 31, 2015; Exclusions: April 1, 2014-March 31, 2015; Exceptions: April 1, 2014-March 31, 2015
Number of Facilities: 425
Effective Sample (denominator after exclusions): 8,166
Level of Analysis: Facility
Patient Characteristics: Not applicable

Comparability of Multiple Data Sources/Methods
N/A

Comparability of Performance Scores when more than one Set of Specifications
N/A- This measure only uses one set of specifications.

Missing Data Analysis and Minimizing Bias
Data Source: Denominator: CDW; Numerator: CDW; Exclusions: CDW; Exceptions: CDW [maintained by CMS]
Dates: Denominator: January 1, 2014-December 31, 2014; Numerator: January 1, 2014-December 31, 2014; Exclusions: January 1, 2014-December 31, 2014; Exceptions: January 1, 2014-December 31, 2014
Number of Facilities: 1,902
Sample (denominator): 64,827
Effective Sample (denominator after exclusions): 13,195
Level of Analysis: Case
Sampled Patient Characteristics: Gender (% Male): 68.5; Mean Age (Years): 62.3 (St. Dev.: 13.8); Race (% Minority): 9.7

1.8 What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example, patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).

We assessed patient-level SDS factors as part of the regression model reported in Section **1b.4**, which provides an overview of disparities in care for patient sub-populations. We based this analysis on SDS variables included in the CDW data:

- Age
- Gender
- Race
- Ethnicity

While an analysis of SDS factors is important in understanding differences in care for patient sub-populations, this measure is a process measure that is neither risk-adjusted nor risk-stratified. We determined that risk adjustment and risk stratification were not appropriate based on the current evidence base and the measure construct. Additional information on this determination is provided in Section **2b4.2**.

2a2. RELIABILITY TESTING
***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)

☒ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)
☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)

N/A

Reliability was calculated in accordance with the methods discussed in *Estimating Reliability and Misclassification in Physician Profiling* (2010). This approach uses a two-level hierarchical linear model (HLM), which is appropriate for testing the reliability of continuous data that have clustered observation (i.e. multiple cases within one facility). Specifically, the testing calculated the variability of facility scores within and between each facility, during the January 2014-December 2014 data collection period, with higher scores indicating greater reliability. The reliability score is a function of the facilities' sample size and score on the measure, and the error variance within and across facilities.

HLM is appropriate for analysis of continuous measures because it takes into account the clustering of cases within a facility and estimates the facility variance as well as the error variance. Reliability for NQF #0290 was calculated at the case level and extreme values for facility score were artificially censored at the 99th percentile (478 minutes). Artificially censoring outlier cases limits the biasing effects of these cases while not rewarding facilities for poor performance. Additionally, facilities with fewer than 11 cases during the January 2014-December 2014 data collection period were omitted in accordance with Hospitals Compare's minimum case count criteria.

See Section **2b.2** for validity testing of data elements.

REFERENCE:

1) Adams J.L., Mehrotra, A., & McGlynn, E.A. Estimating reliability and misclassification in physician profiling. Santa Monica, CA: RAND Corporation. 2010. Retrieved from http://www.rand.org/pubs/technical_reports/TR863.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

N/A

The performance score reliability for NQF #0290 during the January 2014-December 2014 data collection period was 0.78. Reliability is the equivalent of the intraclass correlation coefficient (ICC) and values can range from zero to one. The ICC was calculated using the following equation:

$$\text{ICC} = \frac{variance_{facility}}{variance_{facility} + variance_{error}}$$

The facility variance equaled 4530.2 (95% CI: 4387.3-4677.8) and the error variance equaled 1250.0 (95% CI: 1060.5-1473.3).

REFERENCE:

1) Bartlett, J.W. & Frost, C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. 2008.

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

N/A

Calculated using an HLM model, a reliability score of 0.78 is indicative of strong measure reliability. The result of this test indicates that the measure is able to identify true differences in performance between facilities.

2b2. VALIDITY TESTING

2b2.1. What level of validity testing was conducted? (*may be one or both levels*)

☒ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ **Performance measure score**

☐ **Empirical validity testing**

☒ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)

2b2.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

N/A

The validity of the measure was assessed using both qualitative and quantitative analyses. Data element validity of critical data elements was evaluated by calculating a kappa statistic (for categorical data elements) or Pearson correlation coefficient (for continuous data elements) assessing the level of agreement between facility abstraction and auditor (CDAC) abstraction. Face validity of the performance score was systematically assessed through survey of the EWG.

The test statistics measure interrater reliability and demonstrate the percent agreement between two sources for the same observation, after accounting for agreement by chance. For this test, CDAC is considered to be an authoritative source to which facility abstraction is compared. Test values range from 0.00 to 1.00, where a value of 0.00 indicates zero agreement between two sources and a value of 1.00 indicates complete agreement between two sources. To estimate the statistical significance associated with the test statistics, p-values can be calculated. P-values of less than 0.001 suggest very high levels of statistical significance, and suggest the results are not due to chance. For NQF measure #0290, kappa statistics and Pearson correlation coefficients were used to estimate the level of agreement between the facility's abstraction of critical data elements versus CDAC's abstraction of the critical data elements for the same sample of cases. The p-values associated with these estimates are also reported.

Landis & Koch, 1977 offer the following classification of kappa interpretation:

<0	Poor agreement
0.00–0.20	Slight agreement
0.21–0.40	Fair agreement
0.41–0.60	Moderate agreement
0.61–0.80	Substantial agreement
0.81–1.00	Almost perfect agreement

Similar interpretation is appropriate for Pearson correlation coefficients.

Face validity of the performance score was systematically assessed through survey of the EWG. Five EWG members participated in the data collection. Respondent perspectives include clinicians, management, and healthcare administration. Prior to responding to questions related to performance score face validity, EWG members were provided detailed measure specifications.

The following statements related to performance score face validity were posed to the EWG:

1. Patients who are transferred to another facility for an acute coronary intervention (ACI) can be accurately captured using chart-abstracted data.
2. The measure successfully assesses the timely transfer of AMI patients requiring an ACI.

Responses to statements *1* and *2* in the measure-score face-validity section were collected using a five-point Likert scale: strongly agree, agree, undecided, disagree, strongly disagree, and do not know/not applicable.

REFERENCE:
1) Landis, J. & Koch, G. The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159-174. 1977.

2b2.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

N/A

Results of critical data element validity testing indicate fair to almost perfect levels of agreement between the facilities' abstraction of critical data elements versus CDAC's abstraction of data elements for the same sample of cases. Continued inclusion of a case in performance score calculation is dependent upon earlier data element values reported by the abstractor, and therefore, it is possible that cases may have populated values for critical data elements after they have been excluded from the effective sample. For this reason, test statistics were only calculated for cases that remained in the effective sample at the corresponding step in the algorithm. For example, if a case had a value of "No" for *Initial ECG Interpretation* (thus excluding them from the effective sample) but also had populated values for later data elements, the case would not be considered in any calculations after *Initial ECG Interpretation*. The test statistic and p-value for each critical data element is provided in the table below, as well as the effective sample size used in the calculation.

Data Element	Test Statistic (p-value)	Effective Sample used in Calculation
E/M Code [a]	1.00 (<0.001)	462
Discharge Code [a]	1.00 (<0.001)	462
Patient Age on Outpatient Encounter Date [c]	-	-
ICD-9-CM Principal Diagnosis Code [a]	1.00 (<0.001)	462
Initial ECG Interpretation [a]	0.63 (<0.001)	462
Fibrinolytic Administration [a]	0.93 (<0.001)	141
Transfer for Acute Coronary Intervention [a]	0.47 (<0.001)	113
Discharge Date [b]	1.00 (<0.001)	81
Discharge Time [b]	1.00 (<0.001)	81
Arrival Time [b]	0.99 (<0.001)	81
Measurement Value [c]	-	-
Reason for Not Administering Fibrinolytic Therapy [a]	0.39 (<0.001)	80

a. The test statistic to assess validity for this data element is a Kappa score.
b. The test statistic to assess validity for this data element is a Pearson's correlation.
c. This data element is a calculated value, not an abstracted value.

Results of the face validity assessment indicate that a diverse group of stakeholders support the validity of the measure. Results for each of the questions related to face validity are below.

1. Patients who are transferred to another facility for an ACI can be accurately captured using chart-abstracted data.

Response Option	Response Percentage	Response Count
Strongly Agree	60%	3
Agree	40%	2
Undecided	0%	0

Disagree	0%	0
Strongly Disagree	0%	0
Do Not Know or Not Applicable	0%	0

2. The measure successfully assesses the timely transfer of AMI patients requiring an ACI.

Response Option	Response Percentage	Response Count
Strongly Agree	20%	1
Agree	80%	4
Undecided	0%	0
Disagree	0%	0
Strongly Disagree	0%	0
Do Not Know or Not Applicable	0%	0

2b2.4. What is your interpretation of the results in terms of demonstrating validity? (i*.e., what do the results mean and what are the norms for the test conducted?*)

N/A

Results of the quantitative and qualitative analysis are positive and support the validity of the measure and its calculation. Based on the Landis and Koch classification scale, described in Section **2b2.2**, there was fair to almost perfect agreement between facility and auditor abstraction of data elements. Kappa values ranged from 0.39-1.00 and were statistically significant (Section **2b2.3**). This suggests strong validity for several critical data elements of the measure, as currently specified. Although there are several data elements with test statistics indicating only fair or moderate agreement, these scores may not be representative of all facilities. CMS selects facilities for CDAC validation based on two criteria: 1) facilities for which there is a concern regarding data validity and 2) a random sample of facilities. Based on this selection method, the reported kappa scores for *Transfer for Acute Coronary Intervention*, *Arrival Time*, and *Reason for Not Administering Fibrinolytic Therapy* are likely lower than what would be observed for all facilities nationwide.

Five EWG members, with backgrounds in healthcare administration, management, and clinical expertise in radiology, cardiology, and emergency medicine, provided feedback on the face validity of NQF #0290 through an online survey. All respondents agreed or strongly agreed that patients who are transferred to another facility for an ACI can be accurately captured using chart-abstracted data. Similarly, these five respondents agreed or strongly agreed that NQF #0290 successfully assesses the timely transfer of AMI patients requiring an ACI. The respondents generally support the performance score face validity of NQF #0290.

2b3. EXCLUSIONS ANALYSIS

NA ☐ no exclusions — *skip to section 2b4*

2b3.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)

N/A

We tested measure exclusions and numerator exceptions to determine the prevalence of each exclusion and exception, by facility, and at an aggregate level. The analysis tested measure exclusions and numerator exceptions during the January 2014-December 2014 data collection period. Measure exclusions include all cases meeting one or more criteria listed in Section **1.2c**, above. Numerator exceptions include cases meeting one or more criteria listed in Section **1.2d**, above. To supplement the empirical results, we systematically assessed the face validity of current exclusions and exceptions through survey of the EWG based on responses from five EWG members.

The face validity of exclusions was assessed, using the following statement:

1. The current measure specifications use chart-abstracted data to exclude patients from the denominator population based on the situations listed in the table below[4]. Please evaluate the appropriateness of each of the current exclusion criteria.

Responses to the statement were collected using the following response options: keep, remove, and do not know/not applicable.

2b3.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)

N/A

We examined overall frequencies and proportions of cases excluded for each exclusion/exception criterion, among all sampled cases, for 3,024 facilities submitting cases in 2014. The sampled population included 64,827 cases where a patient (age 18 years or older) presented with an AMI to the ED.

Overall Occurrence and Distribution across Facilities for Measure Exclusions and Exceptions

Data Element	Denominator Exclusion or Numerator Exception?		Overall Occurrence		Distribution Across Facilities (%)		
	Denominator Exclusion	Numerator Exception	N	%	25th	50th	75th
Initial ECG Interpretation	X		38,573	59.5	44.1	60.0	76.7
Fibrinolytic Administration	X		4,600	7.1	0.0	0.0	9.1
Transfer for Acute Coronary Intervention	X		3,386	5.2	0.0	0.0	6.9
Discharge Date		X	25	0.0	0.0	0.0	0.0
Discharge Time		X	104	0.2	0.0	0.0	0.0
Arrival Time		X	7	0.0	0.0	0.0	0.0
Reason for Not Administering Fibrinolytic Therapy		X	4,937	7.6	0.0	0.0	27.4
Total Denominator Exclusions	3 exclusions	-	46,559	71.8	58.3	77.8	100.0
Total Numerator Exceptions	-	4 exceptions	5,073	7.8	0.0	0.0	9.8

[4] Respondents were provided a chart table detailing the situations that exclude a patient from the denominator population of NQF #0290 denominator population.

Data Element	Denominator Exclusion or Numerator Exception?		Overall Occurrence		Distribution Across Facilities (%)		
	Denominator Exclusion	Numerator Exception	N	%	25th	50th	75th
Total Removed from the Denominator or Numerator	7 exceptions and exclusions		51,632	79.6	68.7	89.2	100.0

The EWG provided feedback on the appropriateness of the three denominator exclusions criteria: patients without ST-segment elevation on the ECG performed closed to ED arrival, patients receiving fibrinolytic therapy, and patients who were not transferred to another facility for ACI. Survey respondents unanimously agreed that patients receiving fibrinolytic therapy should be excluded from the effective sample. Four of five respondents indicated that patients without ST-segment elevation on the ECG performed closest to ED arrival and patients who were not transferred to another facility for ACI should be excluded from the effective sample. The survey results indicate that a diverse group of stakeholders generally support the current exclusions for NQF #0290.

2b3.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion*)

N/A

As seen in the table reported in Section **2b3.2** above, the frequency of exclusions/exceptions varied substantially across facilities. Measure exclusion and numerator exception criteria are in alignment with clinical guidelines and also ensure that all cases included in the measure have sufficient information to calculate the performance score. After identification of cases for patients 18 years and older with a principal diagnosis associated with acute myocardial infarction and ST-segment elevation on the ECG closest to ED arrival, measure exclusion and numerator exception criteria are applied:

- *Initial ECG Interpretation* is a denominator exclusion criterion. Cases for patients where *Initial ECG Interpretation* equals "No" are excluded from the effective sample. Overall, 59.5% of cases for patients included in the sample are excluded from the effective sample based on *Initial ECG Interpretation*. There is notable variability in the proportion of cases excluded based on *Initial ECG Interpretation* values across facilities, with an interquartile range of 44.1% to 76.7%. Only cases for patients with ST-segment elevation on the ECG closest to ED arrival are included in the effective sample because in such cases there is a clear clinical need for rapid administration of fibrinolytic therapy.

- *Fibrinolytic Administration* is a denominator exclusion criterion. Cases for patients where *Fibrinolytic Administration* equals "Yes" are excluded from the effective sample. Overall, 7.1% of cases for patients included in the sample are excluded from the effective sample based on *Fibrinolytic Administration*. There is notable variability in the proportion of cases excluded based on *Fibrinolytic Administration* values across facilities, with an interquartile range of 0.0% to 9.1%. Only cases for patients who do not receive fibrinolytic therapy are included in the effective sample because timely transfer is emergent for these cases; whereas, the need for timely transfer would be less emergent for cases that received alternative therapy.

- *Transfer for Acute Coronary Intervention* is a denominator exclusion criterion. Cases for patients where *Transfer for Acute Coronary Intervention* equals "[2] There was documentation the patient was admitted to observation status prior to transfer" or "[3] There was documentation the patient was transferred from this facility's

emergency department to another facility for reasons other than acute coronary intervention, or the specific reason for transfer was unable to be determined from medical record documentation." are excluded from the effective sample. Overall, 5.2% of cases for patients included in the sample are excluded from the effective sample based on *Transfer for Acute Coronary Intervention*. There is some variability in the proportion of cases excluded based on *Transfer for Acute Coronary Intervention* values across facilities, with an interquartile range of 0.0% to 6.9%. Only cases for patients who are transferred for ACI are included in the effective sample because the performance score is dependent upon this criterion.

- *ED Departure Date* is a numerator exception criterion. If *ED Departure Date* is equal to "UTD," the case remains in the effective sample, but a *Measurement Value* is not calculated for that case. Overall, less than 0.1% of cases for patients included in the sample have a "UTD" value for *ED Departure Date*. While there is limited variability in the proportion of excepted cases across facilities, the exception remains important as a "UTD" value for this data element makes it impossible to determine the time to transfer for ACI.

- *ED Discharge Time* is a numerator exception criterion. If *ED Discharge Time* is equal to "UTD," the case remains in the effective sample, but a *Measurement Value* is not calculated for that case. Overall, 0.2% of cases for patients included in sample have a "UTD" value for *Discharge Time*. While there is limited variability in the proportion of excepted cases across facilities, the exception remains important as a "UTD" value for this data element makes it impossible to determine the time to transfer for ACI.

- *Arrival Time* is a numerator exception criterion. If *Arrival Time* is equal to "UTD," the case remains in the effective sample, but a *Measurement Value* is not calculated for that case. Overall, less than 0.1% of cases for patients included in the sample have a "UTD" value for *Arrival Time*. While there is limited variability in the proportion of excepted cases across facilities, the exception remains important as a "UTD" value for this data element makes it impossible to determine the time to transfer for ACI.

- *Reason for Not Administering Fibrinolytic Therapy* is a numerator exception criterion. Cases for patients where *Reason for Not Administering Fibrinolytic Therapy* is equal to "[3] No documented contraindication/reason or Unable to determine (UTD)" are included in the effective sample. Cases for patients where *Reason for Not Administering Fibrinolytic Therapy* is equal to "[1] Documented contraindication/reason" or "[2] Cardiogenic shock" are excluded from the effective sample. Overall, 7.6% of cases for patients included in the sample are included in the effective sample based on *Reason for Not Administering Fibrinolytic Therapy.* There is notable variability in the proportion of cases included based on *Reason for Not Administering Fibrinolytic Therapy* values across facilities, with an interquartile range of 0.0% to 27.4%. Only cases for patients who do not have a clinical or patient-centered reason for not receiving fibrinolytics are included in the effective sample.

Results of the survey of the EWG also support the face validity of the exclusions and exceptions for NQF #0290, and indicate that these exclusions are consistent with prevailing gold standards of care or are necessary to support measure calculation.

2b4. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b4.1. What method of controlling for differences in case mix is used?
☒ **No risk adjustment or stratification**
☐ **Statistical risk model with** Click here to enter number of factors **risk factors**
☐ **Stratification by** Click here to enter number of categories **risk categories**

☐ **Other,** Click here to enter description

2b4.2. If an outcome or resource use measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

N/A

This measure is a process measure for which we provide no risk adjustment or stratification. We determined risk adjustment and stratification were not appropriate based on the measure evidence base and the measure construct. As a process-of-care measure, timely transfer for ACI should not be influenced by SDS factors; rather, adjustment would potentially mask such important inequities in care delivery. Variation across patient populations is reflective of differences in the quality of care provided to the disparate patient population included in the measure's denominator.

During the measure maintenance process, we perform an annual review of the literature, to identify articles and clinical practice guidelines published in the last 12 months, which includes a scan for potential patient subpopulations for which there are differences in the clinical decision to transfer for ACI. This most recent review identified evidence from an observational cohort study that suggest there are gender disparities in transfer time from a non-primary coronary intervention (PCI)-capable hospital to a PCI-capable hospital. A study by D'Onofrio and colleagues demonstrated that women that require PCI are more likely to exceed transfer guidelines for PCI as compared to men (Odds ratio: 2.63; 95% CI: 1.17 – 4.07). This relationship between gender and transfer time remained after adjusting for other sociodemographic, clinical, and organizational factors (D'Onofrio et al., 2015).

However, contrary to the evidence gathered from the literature, stakeholder feedback obtained during the three years of implementation and public reporting has not identified concerns related to SDS factors and need for risk adjustment, supporting the conceptual model upon which the measure is based.

REFERENCES:

1) D'Onofrio G, Safdar B, Lichtman JH, et al. Sex Differences in Reperfusion in Young Patients With ST-Segment–Elevation Myocardial Infarction Results From the VIRGO Study. *Circulation*. 2015;131(15):1324-1332. doi:10.1161/CIRCULATIONAHA.114.012293.

2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or sociodemographic factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*)

N/A

Not applicable - No risk adjustment or stratification was performed.

2b4.4a. What were the statistical results of the analyses used to select risk factors?

N/A

Not applicable - No risk adjustment or stratification was performed.

2b4.4b. Describe the analyses and interpretation resulting in the decision to select SDS factors (e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects)

N/A

Not applicable - No risk adjustment or stratification was performed.

2b4.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)
Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

N/A

Not applicable - No risk adjustment or stratification was performed.

2b4.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

Not applicable - No risk adjustment or stratification was performed.

2b4.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

Not applicable - No risk adjustment or stratification was performed.

2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

Not applicable - No risk adjustment or stratification was performed.

2b4.9. Results of Risk Stratification Analysis:

Not applicable - No risk adjustment or stratification was performed.

2b4.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i*.e., what do the results mean and what are the norms for the test conducted*)

Not applicable - No risk adjustment or stratification was performed.

2b4.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

Not applicable - No risk adjustment or stratification was performed.

2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE
2b5.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

N/A

Differences in performance scores and the mean performance score for facilities meeting public-reporting requirements were tested. For the April 2014-March 2015 data collection period, this included 425 facilities. Additional details of this analysis are provided in Section **2b5.2**.

2b5.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

Q1 2010 Analysis Provider Level
1,050 hospitals submitted 3,164 eligible cases.
Median 68.75 minutes
Min 0 minutes
Max 540 minutes *capped
5th percentile 269 minutes
10th percentile 189 minutes
25th percentile 115 minutes
75th percentile 49 minutes
90th percentile 36 minutes
95th percentile 30 minutes

Distribution of Facility Performance Scores

Mean	Std. Dev.	Min.	10th Percent	Lower Quartile	Median	Upper Quartile	90th Percent	Max.
62.1	38.5	20	33	42	54	69	94	474

2b5.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

N/A

The measure was able to detect facilities with better and worse than average performance. The facility performance scores ranged from 20 minutes to 474 minutes with a median of 54 minutes. Fifty percent of facilities fell within the interquartile range of 42 minutes to 69 minutes. The mean ± SD facility performance score was 62.1 minutes ± 38.5 minutes. By reporting a measure mean (benchmark value), this provides an opportunity for outlying facilities to identify underperformance related to timely transfer for ACI in cases when it is clinically appropriate.

2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS
If only one set of specifications, this section can be skipped.

Note: *This item is directed to measures that are risk-adjusted (with or without SDS factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without SDS factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b6.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

N/A

Not Applicable - this measure uses only one set of specifications.

2b6.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

N/A

Not Applicable - this measure uses only one set of specifications.

2b6.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

N/A

Not Applicable - this measure uses only one set of specifications.

2b7. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b7.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

This measure is calculated using chart-abstracted data. To limit the effects of missing data, abstractors cannot submit a value of "missing" for individual data elements. When they submit a value of "missing" the case is rejected from the abstraction tool. While abstractors cannot submit missing data, they may submit a value of "UTD" for select data elements for which missing information may be more likely; for example, *ED Departure Time*. Cases where a value of "UTD" affects clinical decision making are excluded from the measure. Cases where a value of "UTD" is reflective of poor documentation are included in the denominator but excepted from the numerator. To identify the extent and distribution of cases with a value of "UTD" for a data element, we calculated the frequency of such cases as well as the distribution of cases across eligible facilities. The frequency and distribution of missing data are described in Section **2b3.3.**

2b7.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

The frequency and distribution of missing data are described in Section **2b3.3**. We did not perform statistical analyses of missing data.

2b7.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i.*e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

As described in Section **2b3.3**, the removal of cases from the denominator and/or numerator where an abstractor submits a value of "UTD" are necessary to align with clinical guidelines and enable measure calculation. Additionally, these exclusions/exceptions limit the biasing effects of missing data. Cases where a value of "UTD" affects clinical decision making are excluded from the measure. Cases where a value of "UTD" is reflective of poor documentation are

included in the denominator but excepted from the numerator. Overall, 136 cases of the 13,195 cases in the effective sample (0.2%) have "UTD" value for the three numerator exception criteria, suggesting that such cases have a negligible effect on measure scores. The frequency and distribution of numerator exceptions are discussed in Section **2b3.2**. This exclusion/exception approach penalizes facilities for poor documentation, but does not artificially include cases where transfer for ACI may not reflect appropriate care.

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. **Data Elements Generated as Byproduct of Care Processes.**

Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. **To what extent are the specified data elements available electronically in defined fields?** *(i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields)*

Some data elements are in defined fields in electronic sources

3b.2. **If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources.**

NQF #0290 assesses the time interval from ED arrival to transfer to another facility for ACI for patients with confirmed AMI. Challenges in interpreting and operationalizing the current measure's algorithm make it difficult to respecify this measure for an EHR reporting program since some data elements of the measure currently rely on logic and inferences that abstractors have been trained to interpret. The most significant challenge related to the measure's algorithm involves the Reason for Not Administering Fibrinolytic Therapy data element, which requires a contraindicated/reason for not administering fibrinolytic therapy which is currently located in unstructured fields of the patient record.

Additionally, the potential for e-specification will also require special attention to the Initial ECG Interpretation and Transfer for Acute Coronary Intervention. Abstractors often rely on ECG print-outs and medical notes to determine the presence of ST-segment elevation on an ECG, and if a transfer was specifically for ACI.

Use of EHR data will require vendors to develop mechanisms to capture ECG findings, which are currently used to define the measure's denominator, in a structured field; because of the complicated nature of ECG interpretation, results included in a structured field should come from a review by an eligible provider rather than the ECG machine´s output. Because the reason for patient transfer effectively excludes patients from the measure, these data will also need to be captured and made available for measurement within the EHR workflow.

The AMI and Stroke expert work group (EWG) considers NQF #0290 to be wholly feasible as it is currently specified, but considers e-specification to be moderately feasible. They concur that the key data elements for NQF #0290 are not readily available in a structured format within all EHR systems. In particular, EHR systems may need a new structured field for Initial ECG Interpretation, Reason for Not Administering Fibrinolytic Therapy, and Transfer for Acute Coronary Intervention, which are not perceived to be a standard feature for most systems at this time.

Based on EWG feedback, the availability of the information from the data elements is highly dependent upon the EHR system used in each facility. If they cannot be translated into structured fields, then the data elements must be manually chart abstracted.

3b.3. **If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-**

specific URL.
No feasibility assessment **Attachment:**

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.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 and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.**
IF a PRO-PM, consider implications for both individuals providing PROM data (patients, service recipients, respondents) and those whose performance is being measured.
An online survey of five members of the AMI and Stroke EWG with expertise in cardiology, radiology, emergency medicine, and emergency nursing was conducted to assess the face validity, feasibility, use, and usability of NQF #0290. All participants agreed or strongly agreed that patients who are transferred to another facility for an ACI can be accurately captured using chart abstracted data. Additionally, three of five participants agreed that practical aspects of reporting this chart-abstracted measure do not place undue burden on facilities that collect the data. Those that did not agree indicated the burden will vary between facilities and depend on how well their health record system is constructed. In addition, those that did not agree specifically noted that requiring details about the logistics of a patient transfer as well as the documentation of contraindications may pose a burden on facilities as this is not something that is normally documented in ED workflow.

3c.2. **Describe any fees, licensing, or other requirements to use any aspect of the measure as specified** *(e.g., value/code set, risk model, programming code, algorithm)*.
No fees, licensure, or other requirements are necessary to use this measure; however, CPT codes, descriptions, and other data are copyright 2013 American Medical Association. All rights reserved. CPT is a registered trademark of the American Medical Association. Applicable FARS\DFARS Restrictions Apply to Government Use. Fee schedules, relative value units, conversion factors, and/or related components are not assigned by the AMA, are not part of CPT, and the AMA is not recommending their use. The AMA does not directly or indirectly practice medicine or dispense medical services. The AMA assumes no liability for data contained or not contained herein.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. **Current and Planned Use**
NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Planned	Current Use (for current use provide URL)
	Public Reporting Hospital Outpatient Quality Reporting (HOQR) http://www.medicare.gov/hospitalcompare/search.html Hospital Outpatient Quality Reporting https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1192804531207 Quality Improvement with Benchmarking (external benchmarking to multiple

	organizations) Hospital Outpatient Quality Reporting http://www.medicare.gov/hospitalcompare/search.html Hospital Outpatient Quality Reporting https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1192804531207

4a.1. **For each CURRENT use, checked above, provide:**

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included

Public Reporting:
Name of program and sponsor: CMS HOQR Program
Purpose: The HOQR Program is a pay for quality data reporting program implemented by CMS for outpatient hospital services. In addition to providing hospitals with a financial incentive to report their quality of care measure data, the HOQR Program provides CMS with data to help Medicare beneficiaries make more informed decisions about their health care. Hospital quality of care information gathered through the HOQR Program is publicly available on the Hospital Compare website.
Accountable entities and patients: The publicly reported values (on Hospital Compare) are calculated for all facilities in the United States that meet minimum case count requirements. The number of facilities that met minimum case count criteria during the April 2010-March 2015 data collection periods ranged from 400-425 facilities, annually. The number of facilities meeting minimum case count criteria by year is presented in Section 1b.2. Facilities eligible to report this measure are subject to the Outpatient Prospective Payment System (OPPS) guidelines.

Quality Improvement with Benchmarking (external benchmarking to multiple organizations):
Name of program and sponsor: CMS HOQR Program
Purpose: The HOQR Program is a pay for quality data reporting program implemented by CMS for outpatient hospital services. In addition to providing hospitals with a financial incentive to report their quality of care measure data, the data is publicly reported on the Hospital Compare Website. The data reported on Hospital Compare not only shows the hospital's score on the measure, but also provides state and national averages for the measure. This enables consumers to compare the hospital's performance to other facilities and determine if the facility is an outlier.
Accountable entities and patients: The publicly reported values (on Hospital Compare) are calculated for all facilities in the United States that meet minimum case count requirements. The number of facilities that met minimum case count criteria during the April 2010-March 2015 data collection periods ranged from 400-425 facilities, annually. The number of facilities meeting minimum case count criteria by year is presented in Section 1b.2.Facilities eligible to report this measure are subject to the OPPS guidelines.

4a.2. **If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons?** (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)
This measure is publicly reported.

4a.3. **If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement.** (*Credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.*)
This measure is publicly reported.

4b. Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b.1. **Progress on Improvement. (Not required for initial endorsement unless available.)**
Performance results on this measure (current and over time) should be provided in 1b.2 and 1b.4. Discuss:

- **Progress (trends in performance results, number and percentage of people receiving high-quality healthcare)**
- **Geographic area and number and percentage of accountable entities and patients included**

Summary statistics of performance scores during the April 2010-March 2011 and April 2014-March 2015 data collection periods are

provided in Section 1b.2.

The median value for median time to transfer for ACI for cases where patients have a principal diagnosis associated with AMI and have ST-segment elevation on the ECG closest to ED arrival has declined 3.6% (2 minutes) between the April 2010 and March 2015 data collection periods. Four hundred twenty-one facilities met minimum case counts during the April 2010-March 2011 data collection period, and 425 facilities met minimum case counts during the April 2014-March 2015 data collection period. During the April 2010-March 2011 data collection period, there were 8,008 sampled cases where a patient had a principal diagnosis associated with AMI, had ST-segment elevation on the ECG closest to ED arrival, and was transferred for ACI. Of those patients, the median time to transfer for ACI was 56 minutes. During the April 2014-March 2015 data collection period, there were 8,166 sampled cases where a patient had a principal diagnosis associated with AMI, had ST-segment elevation on the ECG closest to ED arrival, and was transferred for ACI. Of those patients, the median time to transfer for ACI was 54 minutes.

These cases reflect only a subset of the patients eligible for the measure. Dependent upon the facility's total case count, the facility may report all cases or a sample of cases; thus, the number of patients receiving high-quality healthcare as performance on the measure improves is larger than the number of cases captured by the measure.

4b.2. **If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.**
Not applicable, as there is demonstrated improvement in measure performance over time.

4c. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4c.1. **Were any unintended negative consequences to individuals or populations identified during testing; OR has evidence of unintended negative consequences to individuals or populations been reported since implementation? If so, identify the negative unintended consequences and describe how benefits outweigh them or actions taken to mitigate them.**
Measure testing did not identify any unintended consequences. Similarly, no evidence of unintended consequences to individuals or populations has been reported by external stakeholders since its implementation. The potential for unintended consequences will continue to be monitored through an annual review of the literature as well as an ongoing review of stakeholder comments and inquiries.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.
Yes

5.1a. **List of related or competing measures (selected from NQF-endorsed measures)**
0163 : Primary PCI received within 90 minutes of hospital arrival
0288 : Fibrinolytic Therapy Received Within 30 Minutes of ED Arrival

5.1b. **If related or competing measures are not NQF endorsed please indicate measure title and steward.**
Median Time to Fibrinolysis - Centers for Medicare & Medicaid Services (CMS)

5a. Harmonization

The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. **If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):**
Are the measure specifications completely harmonized?
No

5a.2. **If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.**
NQF #0290 and NQF #0288 are both in the HOQR Program, and NQF #0163 is included in the Hospital Inpatient Quality Reporting (HIQR) Program as an electronically specified clinical quality measure (eCQM). While the care settings for the HOQR and HIQR measures differ, all three measures have the same initial patient population – patients with AMI and ST-segment elevation on the ECG performed closest to hospital arrival. While the target populations are the same, the focus of the three measures is different. NQF #0288 focuses on the timely administration of fibrinolytic therapy, NQF# 0290 focuses on the timely transfer of patients who require a PCI, and NQF #0163 focuses on the timely initiation of PCI for a patient who arrives at a PCI-capable hospital. All three measures share a number of key data elements (i.e., Initial ECG Interpretation, Fibrinolytic Administration, and Arrival Time). The specifications for the three measures are generally aligned, where possible.

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. **If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):**
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)
No competing measures that address both the same measure focus and target population as NQF #0290 were identified.

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment:

Contact Information

Co.1 **Measure Steward (Intellectual Property Owner):** Centers for Medicare & Medicaid Services
Co.2 **Point of Contact:** Megan, Hayden, megan.hayden@cms.hhs.gov, 410-786-1970-
Co.3 **Measure Developer if different from Measure Steward:** The Lewin Group
Co.4 **Point of Contact:** Colleen, McKiernan, Colleen.McKiernan@lewin.com, 703-269-5595-

Additional Information

Ad.1 **Workgroup/Expert Panel involved in measure development**
Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
The contractor has convened an EWG, which evaluates and provides feedback on measure-development and maintenance efforts for a set of five AMI and one stroke measure. Specifically, the EWG provides direction and feedback through all phases of project activities, including expansion of the measures to additional CMS quality reporting programs, updates to the current specifications of these six measures, review of quantitative testing results, feedback on qualitative testing questions (i.e., results of EWG member questionnaires), and support for endorsement of the measures by the National Quality Forum (NQF).

The following is a list of the contractor's EWG members:

Joseph P. Drozda, Jr., MD
TEP 2010; Mercy Health, Rep. of American College of Cardiology; Director of Outcomes Research

Mustapha Ezzeddine, MD
University of Minnesota Medical center, Director, Stroke Program

T. Bruce Ferguson, Jr., MD, FACC
TEP 2010; Brody School of Medicine at ECU, Dept. of Cardiovascular Sciences, Professor of Surgery and Physiology

Joseph V. Messer, MD, MACC
TEP 2010; Rush University Medical Center, Rep. of American Medical Association, Professor of Medicine

Cathy Olson, MSN, RN
Emergency Nurses Association (ENA), Institute for Quality, Safety, and Injury Prevention, Director

David Seidenwurm, MD
American Society of Neuroradiology (ASNR); American College of Radiologists (ACR)

Stephen Traub, MD
TEP 2010; Mayo Clinic, Department of Emergency Medicine, Chair

Paul D. Varosy, MD, FACC, FAHA, FHRS
TEP 2010; VA Eastern Colorado Health Care System, Director of Cardiac Electrophysiology

Matt Zavadsky, MS-HPA
National Association of Emergency Medical Technicians (NAEMT)

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 **Year the measure was first released:** 2008

Ad.3 **Month and Year of most recent revision:** 01, 2016

Ad.4 **What is your frequency for review/update of this measure?** Annually

Ad.5 **When is the next scheduled review/update for this measure?** 01, 2017

Ad.6 **Copyright statement:** This measure does not have a copyright.

Ad.7 **Disclaimers:** CPT codes, descriptions, and other data only are copyright 2013 American Medical Association. All rights reserved. CPT is a registered trademark of the American Medical Association. Applicable FARS\DFARS Restrictions Apply to Government Use. Fee schedules, relative value units, conversion factors and/or related components are not assigned by the AMA, are not part of CPT, and the AMA is not recommending their use. The AMA does not directly or indirectly practice medicine or dispense medical services. The AMA assumes no liability for data contained or not contained herein.

Ad.8 **Additional Information/Comments:**

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 2939
De.2. **Measure Title:** Statin Therapy in Patients with Clinical Atherosclerotic Disease
Co.1.1. **Measure Steward:** American College of Cardiology
De.3. **Brief Description of Measure:** Percentage of patients 18-75 year of age with clinical atherosclerotic cardiovascular disease (ASCVD) who were offered moderate-to high-intensity statin.
1b.1. **Developer Rationale:** Use of statin therapy as secondary prevention for patients with clinical atherosclerotic cardiovascular disease can substantially reduce the risk of future cardiovascular events by approximately 30% for patients treated with moderate-intensity and 45% with high-intensity statins (Stone, 2013). Previously, performance measures have focused on managing low-density lipoprotein cholesterol (LDL-C) levels to reduce future atherosclerotic cardiovascular disease (ASCVD) risk. The 2013 American College of Cardiology and American Heart Association guideline release on the treatment of blood cholesterol to reduce ASCVD risk focused its recommendations on statin use and not LDL-C targets due on the overwhelming body of randomized controlled trial evidence showing a clear link between the use of fixed doses of cholesterol-lowering drugs and ASCVD risk reduction. Specifically, the Cholesterol Treatment Trialists provided a comprehensive assessment of the benefits observed with statins. They undertook meta-analyses of individual participant data from 26 RCTs and demonstrated reduction in all-cause mortality, which was largely attributable to significant reductions in deaths due to CAD and other cardiac causes. The majority of studies in the aforementioned report included patients with known ASCVD. Of the 26 RCTs included, 5 trials (39,612 subjects, all of whom had CAD) compared more versus less intensive statin regimens. The trials demonstrated that more intensive regimens produced a highly significant 15% further reduction in major vascular events, driven by reductions in coronary death or nonfatal MI, coronary revascularization, and ischemic stroke. The investigators also found no significant effects observed on deaths due to cancer or other nonvascular causes or on cancer incidence, even at low LDL-C concentrations (Baigent, 2010; Stone, 2013). As such, this measure focuses on the broad clinical ASCVD population and whether optimal treatment (ideally high-intensity statins or moderate-intensity if a contraindication to high-intensity statins exists) is achieved.

References:

1. Baigent C, Blackwell L, Emberson J, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet. 2010;376: 1670–1681.

2. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, Goldberg AC, Gordon D, Levy D, Lloyd-Jones DM, McBride P, Schwartz JS, Shero

3. ST, Smith SC Jr, Watson K, Wilson PWF. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2014;63:2889–2934.

S.4. **Numerator Statement:** Patients in the denominator who have been offered* high-intensity statin† OR have been offered* moderate-intensity statin†.

Definitions:
*A statin is "offered" if it is prescribed or if a patient reason exception for not being prescribed a statin is documented.

†Moderate-intensity and high-intensity statin doses are defined in Table 5 of the 2013 ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adult
http://content.onlinejacc.org/article.aspx?articleid=1879710)

S.7. **Denominator Statement:** All patients 18-75 years of age with clinical ASCVD* who were seen within a 12-month period. This measure is designed to apply to chronic care populations and does not apply to patients in acute care hospitals.

Definition:
*Clinical ASCVD includes acute coronary artery syndromes, a history of MI, stable or unstable angina, coronary or other arterial revascularization, stroke, TIA, and peripheral arterial disease presumed to be of atherosclerotic origin.
S.10. **Denominator Exclusions:** Exceptions: Documentation of medical reason(s) for not prescribing a statin (e.g., allergy, intolerance to statin[s], other medical reasons).

De.1. **Measure Type:** Process
S.23. **Data Source:** Electronic Clinical Data : Registry
S.26. **Level of Analysis:** Clinician : Individual

New Measure -- Preliminary Analysis

Criteria 1: Importance to Measure and Report

1a. Evidence

1a. Evidence. The evidence requirements for a *process or intermediate outcome* measure is that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured.

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☒ **Yes** ☐ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☒ **Yes** ☐ **No**
- **Evidence graded?** ☒ **Yes** ☐ **No**

Evidence Summary:

- The developer provided a diagram demonstrating the steps between the use of fixed doses of cholesterol-lowering drugs as secondary prevention for ASCVD and a reduction in mortality. They note that a more intensive regimen showed a 15% further reduction in major vascular events.
- The developer provided a clinical guideline from the 2013 ACCF/AHA Guideline for the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults with three secondary prevention recommendations:
  - High-intensity statin therapy should be initiated or continued as first-line therapy in women and men <75 years of age who have clinical ASCVD, unless contraindicated. **Class I; Level of Evidence: A**
  - In individuals with clinical ASCVD in whom high-intensity statin therapy would otherwise be used, when high-intensity statin therapy is contraindicated or when characteristics predisposing to statin-associated adverse effects are present, moderate-intensity statin should be used as the second option if tolerated. **Class I; Level of Evidence: A**
  - In individuals with clinical ASCVD >75 years of age, it is reasonable to evaluate the potential for ASCVD risk-reduction benefits and for adverse effects and drug–drug interactions and to consider patient preferences when initiating a moderate- or high-intensity statin. It is reasonable to continue statin therapy in those who are tolerating it. **Class IIa; Level of Evidence: B**
- The developer provided a systematic review of the body of the evidence supporting the use of fixed doses of cholesterol-lowering drugs to reduce ASCVD risk.
- The developer also provided the Quality, Quantity, and Consistency of the body of evidence which included 19 secondary prevention randomized controlled trials and 4 meta-analysis.
- The developer included the potential risk-reduction benefits and adverse effects associated with statin therapy.

Exception to evidence: N/A

Guidance from the Evidence: Process measure with systematic review (Box 3) → Summary of the QQC provided (Box 4) → Systematic review concludes: Quantity: High; Quality: High; Consistency: High (Box 5a) → High

Questions for the Committee:

- *What is the relationship of this measure to patient outcomes?*
- *How strong is the evidence for this relationship?*
- *Is the evidence directly applicable to the process of care being measured?*

Preliminary rating for evidence: ☒ High ☐ Moderate ☐ Low ☐ Insufficient

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- There are no performance scores on the measure as specified. This measure was developed and approved by the ACC and AHA in 2014; however, the developer provided the differences in provider performance from the 2013 and 2014 PINNACLE Registry:

Year	# of providers	# of patients	Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std. Dev.
2013	1701	209770	0.00%	4.30%	16.3%	25.0%	81.4%	20.7%	15.6%
2014	1890	239948	0.00%	8.33%	20.9%	29.0%	87.5%	20.7%	17.3%

- The developer provided a summary of data from the literature that indicates that patients are not receiving optimal statin doses.
  - In a secondary analysis of the TRIUMPH study, 23% of 4,271 patients discharged following an acute MI were on maximal statin therapy with variability across hospitals (Arnold, 2014).
  - In a study of the GWTG Registry, 38.3% of 65,396 patient with ACS who were discharged with lipid-lowering agents were discharged with intensive lipid-lowering therapy (Javed, 2011).
  - Another study examined prescribing practices of statins and found a variation, specifically by diagnosis and age, in titration of high-intensity doses of statins for CAD and PAD patients . (Jeevanantham, 2015).
- The developer included additional publications that demonstrated statin therapy continues to be underutilized in the broad ASCVD population.

Disparities:

- The developer noted that there is limited published data on potential disparities in care for ASCVD patients receiving statin therapy:
  - Blacks (47%) were less likely than whites (52%) to report receiving statins (Lipworth, 2014).
  - Patients at a higher risk for cardiovascular disease: blacks reported statin use 38% and whites 50% (Qato, 2010).
  - 47% of black men were prescribed a statin while close to 60% of white men were prescribed lipid-lowering agents (Massing, 2004).
  - Uninsured patients were 6% less likely to receive lipid-lowering therapy (Smolderen, 2013).

Questions for the Committee:

- *Does the data from the literature presented demonstrate that ASCVD patients are not receiving optimal (high- or moderate-intensity) statin therapy? Is a national performance measure warranted?*
- *Are you aware of evidence that disparities exist in this area of healthcare?*

Preliminary rating for opportunity for improvement: ☒ High ☐ Moderate ☐ Low ☐ Insufficient

Committee pre-evaluation comments

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence to Support Measure Focus

Comments: **This is a process measure and the evidence is high

**This is a process measure where there is evidence that adherence to the measure is associated with improved outcomes. that

association is documented in the references.

**There is extensive RCT trial data and meta-analyses showing CV event lowering in patients with ASCVD who receive moderate to high intensity statins. Very strong data to support recommendation

The measure is directly applicable to the process of care being measured.

1b. Performance Gap

Comments: **There is a remarkably large performance gap.

T**here are a number of guidelines (ACC, National Lipid Association, and others that support the guidelines; however ACC points out that there are no performance scores. However there are studies suggesting that patients may not be receiving the proper care.

**There is a large gap and this performance measurement would hopefully help improve this gap in optimal care with the use of moderate to high intensity statins in patients with ASCVD. There are health disparities when looking at race (black vs nonblack) and insured vs uninsured in the literature relative to preventative CV prescribing.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

2a1. Reliability Specifications

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented.

Data source(s): electronic clinical registry data. This is not an eMeasure.

Specifications:

- The level of analysis is at the clinician-level.
- The numerator includes patients who have been offered a high-intensity statin or have been offered a moderate-intensity statin
    - A statin has been "offered" if it has been prescribed or if a patient reason exception for not being prescribed a statin is documented
    - Moderate- and high-intensity statins are defined in the ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults available at http://content.onlinejacc.org/article.aspx?articleid=1879710
- The denominator includes all patients ages 18-75 with clinical ASCVD who were seen within a 12-month period (for chronic care population and not acute care hospitals)
    - Clinical ASCVD includes acute coronary artery syndromes, a history of MI, stable or unstable angina, coronary or other arterial revascularization, stroke, TIA, and peripheral arterial disease presumed to be of atherosclerotic origin.
- The denominator exception includes documentation of medical reason(s) for not prescribing a statin (e.g., allergy, intolerance to statin[s], other medical reasons).
- The calculation algorithm is included.
- Details for handling missing data are provided.
- The collection instrument and data dictionary are included in the attachment titled ASCVDStatinAppendixA1.pdf.

Questions for the Committee :

- *Are all the data elements clearly defined?*
- *Is the logic or calculation algorithm clear?*
- *Is it likely this measure can be consistently implemented?*

2a2. Reliability Testing attachment

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers.

SUMMARY OF TESTING

Reliability testing level ☒ **Measure score** ☐ **Data element** ☐ **Both**

Reliability testing performed with the data source and level of analysis indicated for this measure ☒ **Yes** ☐ **No**

Method(s) of reliability testing:

- The final dataset included 2013 and 2014 PINNACLE Registry data for patients with ASCVD:

	2013	2014
Providers w/min. eligible pts. (10)	1,701	1,890
Average # eligible pts. for providers	123.2	126.9
Pt. to provider range	10 - 1,427	10 - 879
# of practices	121	136
Total # of patients	209,770	239,948

- The developers described the validation process used to obtain the final dataset:
  - 2,503,852 patients were treated by 206 practices in 2014. Of those, 68 practices (678,475 patients) were excluded because their EHR did not transmit the data on medication dose.
  - An additional 841,376 patients were excluded due to missing comorbidity data needed to determine the presence of ASCVD.
  - Another 205,345 patients were excluded because the developer was unable to define the statin dose or there was a medical exclusion for statins.
  - Lastly, 16,188 patients and 2 additional practices were excluded because the providers treated <10 eligible patients.
- The developers used a beta-binomial model to assess the signal-to-noise ratio. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in performance. The higher the reliability score, the greater is the confidence with which one can distinguish the performance of one provider from another. This is an appropriate test for measure score reliability. A reliability of 0.70 is generally considered a minimum threshold for reliability.

Results of reliability testing:

- The developer provided the following reliability scores from 2013 and 2014:

2013

Description	Number of Patients	Signal-to-Noise Ratio
Minimum	10	0.985
25th percentile	53	0.991
50th percentile	94	0.993
75th percentile	161	0.996
Average (mean)	124	0.995

2014

Description	Number of Patients	Signal-to-Noise Ratio
Minimum	10	0.986
25th percentile	57	0.992
50th percentile	102	0.994
75th percentile	168	0.996
Average (mean)	127	0.995

Guidance from the Reliability Algorithm : Precise specifications (Box 1)→ Empirical testing as specified (Box 2) → Reliability testing was conducted with performance measure scores for each measure entity (Box 4)→ Method for

assessing proportion of variability due to real differences described (Box 5) → There is high confidence in the performance measure score reliability (Box 6) → High

Questions for the Committee:

- *Is the test sample adequate to generalize for widespread implementation?*
- *Do the results demonstrate sufficient reliability so that differences in provider performance can be identified?*

Preliminary rating for reliability: ☒ High ☐ Moderate ☐ Low ☐ Insufficient

2b. Validity

2b1. Validity: Specifications

2b1. Validity Specifications. This section should determine if the measure specifications are consistent with the evidence.

Specifications consistent with evidence in 1a. ☒ Yes ☐ Somewhat ☐ No
Specification not completely consistent with evidence

Question for the Committee:

- *Are the specifications consistent with the evidence?*

2b2. Validity testing

2b2. Validity Testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

SUMMARY OF TESTING

Validity testing level ☒ Measure score ☐ Data element testing against a gold standard ☐ Both

Method of validity testing of the measure score:

- **☒ Face validity only**
- **☐ Empirical validity testing of the measure score**

Validity testing method:

- Face validity of the measure score was systematically assessed by members of two existing committees from the ACC and AHA who were not involved in the development of the measure.
- The developer did not identify the committee members or provide the total number of committee members assessing face validity.

Validity testing results:

- The developer noted that 8 committee members completed the survey and 100% of the respondents either agreed or strongly agreed that this measure can accurately distinguish good and poor quality.

Questions for the Committee:

- *Do the results demonstrate sufficient validity so that conclusions about quality can be made?*
- *Do you agree that the score from this measure as specified is an indicator of quality?*

2b3-2b7. Threats to Validity

2b3. Exclusions:

- The only exclusion for this measure is documentation of a medical reason(s) for not prescribing a statin (e.g., allergy, intolerance to statin[s], other medical reasons). No patients in either the 2013 or 2014 sample data, had such a contraindication. Although, the developer stated, there were a few patients identified as having patient-centered reasons for not receiving statins, these were not considered exclusions from this performance measure.
- The developer raised one potential opportunity for misclassification of patients who cannot tolerate higher doses of statins and can only be treated with low-dose statins may be misclassified as not meeting the

numerator. This would result in slightly lower performance rate among some physicians.

Questions for the Committee:

- *Are the exclusions consistent with the evidence?*
- *Are any patients or patient groups inappropriately excluded from the measure?*
- *Are the exclusions/exceptions of sufficient frequency and variation across providers to be needed (and outweigh the data collection burden)?*

2b4. Risk adjustment: **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

2b5. Meaningful difference *(can statistically significant and clinically/practically meaningful differences in performance measure scores can be identified)*:

- The developer provided the differences in provider performance from the 2013 and 2014 PINNACLE Registry:

Year	# of providers	# of patients	Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std. Dev.
2013	1701	209770	0.00%	4.30%	16.3%	25.0%	81.4%	20.7%	15.6%
2014	1890	239948	0.00%	8.33%	20.9%	29.0%	87.5%	20.7%	17.3%

- The developer noted that while there may be some opportunity to improve performance by better documentation of medical reasons for not prescribing a statin (such as intolerance), this is unlikely to account for the large differences in performance scores.
- The developer also provided the variation in provider performance based on sex, age, race, and insurance status for 2013 and 2014.
- Overall, a large amount of variability was noted among providers. The developer assessed the Median Rate Ratio (MRR) to determine the average likelihood that a patient with ASCVD would be treated with moderate/high-dose statins by one provider as compared to another.
- For 2013, the MRR was 3.03, which indicated a greater than 3-fold average likelihood a patient would be treated with moderate/high-dose statins by one provider as compared to another. The MRR for 2014 was slightly smaller, 2.41, or an almost 2.5-fold average likelihood.

Question for the Committee:

- *Does this measure identify statistically significant and meaningful differences in performance among providers?*

2b6. Comparability of data sources/methods:

- Measure is not specified for more than one data source; comparability of data sources is not needed.

2b7. Missing Data

- The developer described the analytic method used for handling missing data in Section 1.2 and determined most missing data were due to administrative reasons and not likely to introduce bias into the performance results.
- The developer provided the rate of missing data among providers with statin dose obtained for some patients but not others for 2013 and 2014:

Year	Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std. Dev.
2013	0.000	0.803	0.862	0.962	1.000	0.159	0.145
2014	0.050	0.715	0.806	0.926	1.000	0.211	0.158

Guidance from Validity Algorithm: Specifications consistent with evidence (Box 1) → Potential threats to validity assessed (Box2) → Empirical validity testing not conducted (Box 3)→Face validity systematically assessed by expert panel (Box 4) → Agreement by panel that performance measure score distinguishes quality (Box 5)→ Moderate (highest eligible rating is MODERATE)

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2d)

2a1. & 2b1. Specifications

Comments: **The specifications are clear

**There is evidence of reliability of the measure IF it is drawn from the PINNACLE registry. What is not presented is data from "real world experience" outside of the PINNACLE database to demonstrate the reliability of the measure. Ideally, the sponsor would bring in claims data, EHR data, etc. to demonstrate the reliability of the measure outside of PINNACLE

**The data elements are clearly defined and the measure can be consistently implemented. The limitation is that about 25% of EHRs do not have the capability to report dose of statin which is a limitation.

**There are no inconsistencies

**There has been more recent clinical guidelines that recognize the role of PCSK9 in Familial Hypercholesterolemia (https://www1.ghc.org/static/pdf/public/guidelines/ascvd-secondary.pdf) that may impact the validity of the measure as presented

**The specifics of what the measure is collecting relative to dose of statin prescribed in consistent with the evidence showing benefit in patients with ASCVD prescribed moderate to high intensity statins.

2a2. Reliability Testing

Comments: **Reliability is very high

**Yes, the reliability testing was appropriate for the specific PINNACLE registry use

**Yes, this was performed and there was high reliability with an average signal-to-noise score of 0.99 (> 0.70 is the usual threshold).

2b2. Validity Testing

Comments: **Validity is face validiity. This is not concerning.

**There is evidence for this measure with face validity

**Face validity from the committee members for ACC/AHA and 100% agreement.

2b3. Exclusions Analysis

2b4. Risk Adjustment/Stratification for Outcome or Resource Use Measures

2b5. Identification of Statistically Significant & Meaningful Differences In Performance

2b6. Comparability of Performance Scores When More Than One Set of Specifications

2b7. Missing Data Analysis and Minimizing Bias

Comments: **Secction 2b3 says that there are no exclusions for patient reasons--really?

There is potential for misclassification if the patient is on a low dose statin for intolerance reasons.

Risk adjustment was not done.

**That the PINNACLE data shows a reasonably significant difference among prescribers on a year over year basis, but variability in the Median Rate Ratio (MRR), does raise further the question of potential real world experience. Much of the issue seems to revolve around the documentation of reason for exclusion. This may potentially weaken the strength of this measure.

**Exclusions are appropriate, but there were no patients that met exclusion criteria in the 2013 and 2014 cohorts evaluated. No risk adjustment. There is meaningful differences when evaluating provider to provider with a 3 fold difference in likelihood to receive a moderate to high intensity statin when looking across providerss. Data is not specified for more than one data source. Missing data was mainly for administrative reasons and has a low likelihood to introduce bias into the performance measure.

Criterion 3. Feasibility

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- The developer requests licensing agreement prior to use of this measure but does not provide a description of fees associated with the PINNACLE Registry or whether participation in the registry is limited to cardiologists.
- All data elements are in electronic sources.
- Feasibility issues identified in Section 1.2 included:
    - EHRs that were unable to transmit 'medication dose' from the physician practice to the registry, therefore, were excluded from the sample – this will be addressed in the future by remapping their data extraction.
    - For those practices with the ability to transmit their data electronically, comorbidity data needed to determine the presence of ASCVD and statin dose were some of the missing data elements.

Questions for the Committee:

- *Are the required data elements routinely generated and used during care delivery?*
- *Are the required data elements available in electronic form, e.g., EHR or other electronic sources?*
- *Is the data collection strategy ready to be put into operational use?*

Preliminary rating for feasibility: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee pre-evaluation comments

Criteria 3: Feasibility

3a. Byproduct of Care Processes

3b. Electronic Sources

3c. Data Collection Strategy

Comments: **It appears that some EHRs are not able to transmit the required elements.
**The value of this measure is enhanced if it is not restricted to the PINNACLE registry. The majority of people with ASCVD are managed by non-cardiologists. For that reason the measure should be feasible through data collection methods beyond the ACC PINNACLE Database. I would be happier if the feasibility were demonstrated using claims or EHR data
**There is some concern about feasibility as 25% of EHRs were excluded as they did not report the dose of statin. Most required elements are in all EHRs.

Criterion 4: Usability and Use

4. Usability and Use evaluate the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

Current uses of the measure

Publicly reported? ☐ **Yes** ☒ **No**

Current use in an accountability program? ☐ **Yes** ☒ **No**
OR
Planned use in an accountability program? ☐ **Yes** ☒ **No**

Accountability program details:

- The developer did not provide a credible plan for use in an accountability program within 3 years of initial endorsement. NQF criteria states that for measures not in use at the time of initial endorsement, a credible plan for implementation within the specified timeframes is provided.
- The measure is currently used for quality improvement in the PINNACLE Registry. PINNACLE assists practices in understanding and improving care through the production and distribution of quarterly performance reports. These reports, covering all valid patient encounters, detail adherence to 22 cardiovascular clinical measures at the physician, location, and practice levels across coronary artery disease, hypertension, heart failure and atrial fibrillation.
- The developer stated that they would welcome the implementation of this measure in emerging applications such as accountable care organizations (ACO), Medicare Advantage insurance plans or health plans selling on the new insurance marketplace.

Improvement results:

- The developer provided the mean performance rates and number of participating providers from the PINNACLE Registry from 2013 to 2014.

Potential harms:

- The developer is not aware of any unintended consequences at this time, but continuously monitor for them.

***Questions for the Committee**:*

- *How can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

Preliminary rating for usability and use: ☐ **High** ☐ **Moderate** ☒ **Low** ☐ **Insufficient**

Rationale: Use is limited to a closed registry for quality improvement purposes.

Committee pre-evaluation comments

Criteria 4: Usability and Use

4a. Accountability and Transparency

4b. Improvement

4c. Unintended Consequences

Comments: **The measure is currently in use only for internal QI.

However, the measure is quite simple and should be accessible for other purposes."

**This measure is not currently in use in any public reporting or accountability program. At the same time there are a number of other, similar measures which enjoy greater degrees of support.

**This is not publicly reported currently. Potentially, public reporting could improve the percentage meeting this measure.

Criterion 5: Related and Competing Measures

Related or competing measures:

- 0074 : Chronic Stable Coronary Artery Disease: Lipid Control
- 0118 : Anti-Lipid Treatment Discharge
- 0439 : STK-06: Discharged on Statin Medication
- 1519 : Statin Therapy at Discharge after Lower Extremity Bypass (LEB)
- 696: STS CABG Composite Score (STS)
- 964: Therapy with aspirin, P2Y12 inhibitor, and statin at discharge following PCI in eligible patients (ACC)
- 2452: Percutaneous Coronary Intervention (PCI): Post-procedural Optimal Medical Therapy (ACC)

Harmonization:

- This new measure, #2939, focuses on optimal treatment of statins. Most measures on statin therapy that are NQF-endorsed address subsets of the patients included in this broad measure and do not yet reflect the updated recommendations and/or are intended to be used in a different setting, level of analysis or different data source.

Pre-meeting public and member comments

Comment by: Denise Cannaday, CCO

Organization: SMT Inc.

Comments: #5656, #5657, #5658, and #5659

As the CDC has pointed out, cardiovascular disease accounts for more than one third of all US deaths, and in 2010, the total costs of cardiovascular diseases were estimated to be $444 billion. ([1]http://www.cdc.gov/chronicdisease/resources/publications/AAG/dhdsp.htm) We applaud NQF for making ASCVD a priority and establishing measures to promote appropriate focus on patients who have an ASCVD event. The statin measures will increase the number of patients having cardiovascular care, and this is an important objective for the patient, society, and overall healthcare. This measure as drafted, however, may have some unintended consequences that should be addressed before endorsement.

Primarily, this measure only looks at initiation of treatment and not whether the patient achieved a desired therapeutic response and maintained it over time. Cardiovascular events will continue to be a number one killer if the focus does

not encompass the ongoing treatment of the patient, and modification of their care as needed. Second, the measure only addresses moderate and high intensity statin use for ASCVD patients, which fails to include other high risk groups such as Familial Hypercholesterolemia (FH) patients or diabetics, and does not support patient centric prescribing based upon response to medication, side-effects or available options outside the traditional statins.

The AHA/ACC 2013 guidelines address 4 steps required to treat to an LDL-C goal in order to achieve favorable impact on CV outcomes: Step 1) recognition of risk and LDL-C baseline measurement, Step 2) initiate appropriate treatment based on the LDL-C level and the patient subpopulation and level of ASCVD risk, Step 3) monitor patient response and intensify treatment based on response to achieve desired therapeutic response (Figure 5, Stone, 2013: 3024-3025) and 4) once the therapeutic response is achieved, maintaining it over time. To achieve the goal of treating hyperlipidemia and to reduce overall ASCVD risk, there needs to be a "family" of measures that address all the needed steps.

The first step of diagnosing and initiating treatment is addressed in this process measure but it is not sufficient to achieve the desired clinical outcome for the patient. It does not address Step 3, whether the patient achieved the desired therapeutic response or Step 4 – maintaining the therapeutic response over time. The recommendation is to build the measure to include monitoring lipid levels and adjusting treatment to achieve reduction of CV risk and improve outcomes.

Step 3 comes with the highest level of evidence and recommendation in the AHA/ACC guidelines (Grade A, Strong Recommendation). The evidence cited in the main document and ASCVD Statin Evidence Form supports the importance of this step in clinical practice and documents it as a gap in care. Despite treatment with lipid-lowering therapy, many patients do not achieve optimal LDL-C control and may remain at risk for CV events. (Waters 2009:28-34) This suboptimal response may be due to many factors, including: efficacy limitations, intolerance to lipid-lowering therapy, and treatment non-adherence, including failure to fill the prescription. (Karalis, 2012: doi:10.1155/2012/ 861924) These factors are recognized in the assessment of response and treatment decision-making, as reflected in the treatment algorithm (Figure 5). The 2013 ACC/AHA guidelines state that the relative reduction in ASCVD risk "from statin therapy is related to the degree by which the LDL-C is lowered." (P2901) In addition, there are a number of Evidence Statements in Appendix 4 that reinforce the relationship between reduction in risk linked not just to the dosing but the percent reduction in LDL-C that was achieved. ES 33 states that "In adults 40 to 75 years of age with diabetes and _>1 risk factor, fixed moderate-dose statin therapy that achieved a mean LDL-C-C of 72 mg/dL reduced the RR for CVD by 37% (in this trial, LDL-C-C was reduced by 46 mg/dL or 39%)."

An alternative is not to focus on process, the dosing of a drug initially (Step 2) or on monitoring and intensification (Step 3), but to focus on an intermediate outcome measure which addresses Step 4, whether the patient achieved the desired level of response, reduction of their LDL-C to an acceptable level, which could be an absolute LDL-C level or to the lowest LDL-C level possible for the individual patient. While there is concern with this approach, the LDL-C level is what influences the patients' CV risk. While we recognize the rationale for not recommending use of absolute target LDL-C levels, the current reality is that patients are familiar with having an LDL-C level to guide decisions about lifestyle management and the adequacy of the current treatment regimen and need for adjustment/change. LDL-C reduction provides a clear goal to support patient engagement and patient engagement is a critical element in improving healthcare. Measures need to be easily understood, actionable and provide the transparency necessary for patients and payers to select providers.

Measures need to be crafted to capture relevant evidence needed to provide optimal care. Prescribing of a statin only measure fails to meet the need and has already been shown to be ineffective at capturing the necessary data to improve the quality of care. CMS proposed to retire the statin prescription upon discharge after AMI since performance amongst hospitals was so high and unvarying that meaningful distinctions and improvements could no longer be made.

There is an inherent deficiency in a statin only measure that lacks clinical utility, yet adds a data collection and reporting burden on providers. US physician practices spend more than $15.4 billion annually to report quality measures (Lawrence P. Casalino, David Gans, Rachel Weber, Meagan Cea, Amber Tuchovsky, Tara F. Bishop, Yesenia Miranda, Brittany A. Frankel, Kristina B. Ziehler, Meghan M. Wong and Todd B. Evensong) and Health Affairs found that 12.5 hours

of physician and staff time per physician per week – was spent on "entering information onto the medical record ONLY for the purpose of reporting for quality measures from external entities. Reporting on an outcome, or an intermediate outcome such as LDL-Cs, does serve a clinical purpose and becomes a worthwhile task for provider and patient to improve the quality of care.

Second, our concern is it also fails to address the patient's response to treatment and the consideration of non-statins to achieve the desired response, such as the new PCSK9s. Exclusion of this drug class fails to ensure high quality care to this subset of high risk patients, a quality of care that is possible through individualized care using drugs currently available, which may include a combination of statins and non-statins. The evidence is there to support this action, including the 2016 Pathway for non-statins (Lloyd-Jones 2016) This may impact up to 40-60% of the patients. Maddox et al reported that ~50% of patients were on a statin alone and ~ 30% of patients in all subpopulations are on combination therapy. (Maddox 2014 JACC; 2183-2192). For many, the combination of a statin with another drug (ezetimibe) is needed to achieve the desired reduction. (Karalis 2012) The study IMProved Reduction of Outcomes: Vytorin Efficacy International Trial (IMPROVE-IT) showed a 2% reduction in CV events with Vytorin (ezetimibe/simvastatin) vs. simvastatin, and the results are consistent with the established relationship between absolute reductions in LDL-C and reductions in major CV events (Baigent, 2010: 167-81). Murphy et al established that treatment with ezetimibe plus simvastatin resulted in a 9% reduction in total events of which 56% were first events and 44% were subsequent events. Reduction in total events was driven by reductions in MI and stroke. (Murphy 2016)

Sabatine et al went further and found that there were longer term and continued benefits from the combination therapies from reduced LDL-Cs. Sabatine et al reported the rate of cardiovascular events at 1 year was reduced from 2.18% in the standard-therapy group to 0.95% in the evolocumab group. (Sabatine 2015) Robinson et al reported a reduction in mortality as well as nonfatal events from 3.3% in the placebo group to 1.7% in the intervention group. (Robinson 2015) Both studies treated with PCSK9 to lower LDL cholesterol levels and showed approximately 50% reductions in composite CV events at 12 – 18 months. Compared with less intensive regimens that didn't adjust treatment protocols to the patient, more intensive and patient appropriate care produced a highly significant 15% (95% Cl 11-18; $p<0.0001$) further reduction in major vascular events, consisting of separately significant reductions in coronary revascularization of 19% (95% Cl 15-24; $p<0.0001$), and in ischemic stroke of 16% (95% Cl 5-26; $p= 0.005$.

This has particular import to those with FH for whom LDL-C testing, diagnosis, diet and exercise compliance, dose adjustment, switching or combination and patient engagement is key. The 20-fold increased lifetime risk of ASCVD for those with FH can be reduced to that of the general population with early identification and aggressive management which includes initiation of high-intensity statin therapy, follow-up lipid values, compliance to lifestyle and medical therapy and use of non-statins drugs when indicated based on the individual's response to statins alone. (Knowles AmJCard, 2015; 481-4). Many of these patients are already not receiving recommended therapy with non-statins because of physician concern/inertia and payer barriers. Including non-statin use, such as PCSK9s, in the measure would support the physician in providing optimal care to this subset or other high-risk patients.

In summary, we support this measure in general but are concerned that it is possible to pass the measure by ordering a statin without regard to the patient's response to care. We believe that an additional intermediate outcome element added to the measure or a companion intermediate outcome measure that addresses whether the patient achieved a desired therapeutic response as demonstrated by one of the 'indicators of anticipated therapeutic response" identified in the 2013 Guidelines, e.g. percent reduction or LDL-C level, are needed to ensure that the patient's CV risk will be reduced and to ensure that the focus remains on the overall outcome of care of the patient and not just initiation of therapy (Stone, Table 10).

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** Click here to enter NQF number

Measure Title Statin Therapy in Patients with Clinical Atherosclerotic Disease

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: 4/27/2016

Instructions

- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- Respond to all questions as instructed with answers immediately following the question. All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 10 pages (*incudes questions/instructions*; minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the evidence for this measure meets NQF's evaluation criteria.

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- Health outcome: [3] a rationale supports the relationship of the health outcome to processes or structures of care. Applies to patient-reported outcomes (PRO), including health-related quality of life/functional status, symptom/symptom burden, experience with care, health-related behavior.
- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- Efficiency: [6] evidence not required for the resource use component.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the U.S. Preventive Services Task Force (USPSTF) grading definitions and methods, or Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines.

5. 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 multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's Measurement Framework: Evaluating Efficiency Across Episodes of Care; AQA Principles of Efficiency Measures).

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☐ Health outcome: Click here to name the health outcome

☐ Patient-reported outcome (PRO): Click here to name the PRO

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors

☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

☒ Process: Statin therapy for patients with clinical atherosclerotic cardiovascular disease (ASCVD)

☐ Structure: Click here to name the structure

☐ Other: Click here to name what is being measured

HEALTH OUTCOME/PRO PERFORMANCE MEASURE *If not a health outcome or PRO, skip to 1a.3*

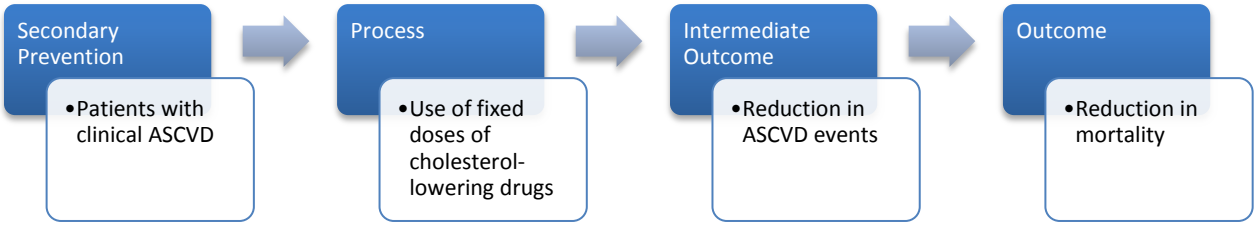
1a.2. Briefly state or diagram the path between the health outcome (or PRO) and the healthcare structures, processes, interventions, or services that influence it.

1a.2.1. State the rationale supporting the relationship between the health outcome (or PRO) to at least one healthcare structure, process, intervention, or service (*i.e., influence on outcome/PRO*).

Note: For health outcome/PRO performance measures, no further information is required; however, you may provide evidence for any of the structures, processes, interventions, or service identified above.

INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURE

1a.3. Briefly state or diagram the path between structure, process, intermediate outcome, and health outcomes. Include all the steps between the measure focus and the health outcome.



Secondary prevention of cardiovascular events in patients with clinical atherosclerotic cardiovascular disease is linked to high- or moderate-intensity statin therapy. Specifically, the Cholesterol Treatment Trialists provided a comprehensive assessment of the benefits observed with statins. They undertook meta-analyses of individual participant data from 26 RCTs and demonstrated reduction in all-cause mortality, which was largely attributable to significant reductions in deaths due to CAD and other cardiac causes. The majority of studies in the aforementioned report included patients with known ASCVD. Of the 26 RCTs included, 5 trials (39,612 subjects, all of whom had CAD) compared more versus less intensive statin regimens. The trials demonstrated that more intensive regimens produced a highly significant 15% further reduction in major vascular events, driven by reductions in coronary death or nonfatal MI, coronary revascularization, and ischemic stroke. The investigators also found no significant effects observed on deaths due to cancer or other nonvascular causes or on cancer incidence, even at low LDL-C concentrations.

Baigent C, Blackwell L, Emberson J, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet. 2010;376: 1670–1681.

1a.3.1. What is the source of the systematic review of the body of evidence that supports the performance measure?

☒ Clinical Practice Guideline recommendation – ***complete sections 1a.4, and 1a.7***

☐ US Preventive Services Task Force Recommendation – ***complete sections 1a.5 and 1a.7***

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*) – ***complete sections 1a.6 and 1a.7***

☐ Other – ***complete section 1a.8***

Please complete the sections indicated above for the source of evidence. You may skip the sections that do not apply.

_____

1a.4. CLINICAL PRACTICE GUIDELINE RECOMMENDATION

1a.4.1. Guideline citation (*including date*) and **URL for guideline** (*if available online*):

Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, Goldberg AC, Gordon D, Levy D, Lloyd-Jones DM, McBride P, Schwartz JS, Shero ST, Smith SC Jr, Watson K, Wilson PWF. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2014;63:2889–2934.

http://content.onlinejacc.org/article.aspx?articleid=1879710

Full report supplement of this guideline (cited in 1a.7) can be accessed at:
http://jaccjacc.cardiosource.com/acc_documents/2013_FPR_S5_Blood_Cholesterol.pdf

1a.4.2. Identify guideline recommendation number and/or page number and **quote verbatim, the specific guideline recommendation**.

2013 ACCF/AHA Guideline for the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults

Secondary Prevention Recommendations (e2900):

1. High-intensity statin therapy should be initiated or continued as first-line therapy in women and men ≤75 years of age who have clinical ASCVD, unless contraindicated. **Class I; Level of Evidence: A**
2. In individuals with clinical ASCVD in whom high-intensity statin therapy would otherwise be used, when high-intensity statin therapy is contraindicated or when characteristics predisposing to statin-associated adverse effects are present, moderate-intensity statin should be used as the second option if tolerated. **Class I; Level of Evidence: A**
3. In individuals with clinical ASCVD >75 years of age, it is reasonable to evaluate the potential for ASCVD risk-reduction benefits and for adverse effects and drug–drug interactions and to consider patient preferences when initiating a moderate- or high-intensity statin. It is reasonable to continue statin therapy in those who are tolerating it. **Class IIa; Level of Evidence: B**

"Clinical ASCVD includes acute coronary syndromes, history of MI, stable or unstable angina, coronary or other arterial revascularization, stroke, TIA, or peripheral arterial disease presumed to be of atherosclerotic origin."

1a.4.3. Grade assigned to the quoted recommendation with definition of the grade:

Guideline Statement #	Class of Recommendation/Level of Evidence (for definitions see 1a.4.4 below)

(see 1a.4.2 above)	
1	Class Ia
2	Class Ia
3	Class IIa B

1a.4.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: If separate grades for the strength of the evidence, report them in section 1a.7.*)

Class of Recommendation (COR) is an estimate of the size of the treatment effect considering risks versus benefits in addition to evidence and/or agreement that a given treatment or procedure is or is not useful/effective or in some situations may cause harm.

Class I: Procedure/Treatment should be performed/administered

Class IIa: It is reasonable to perform procedure/administer treatment

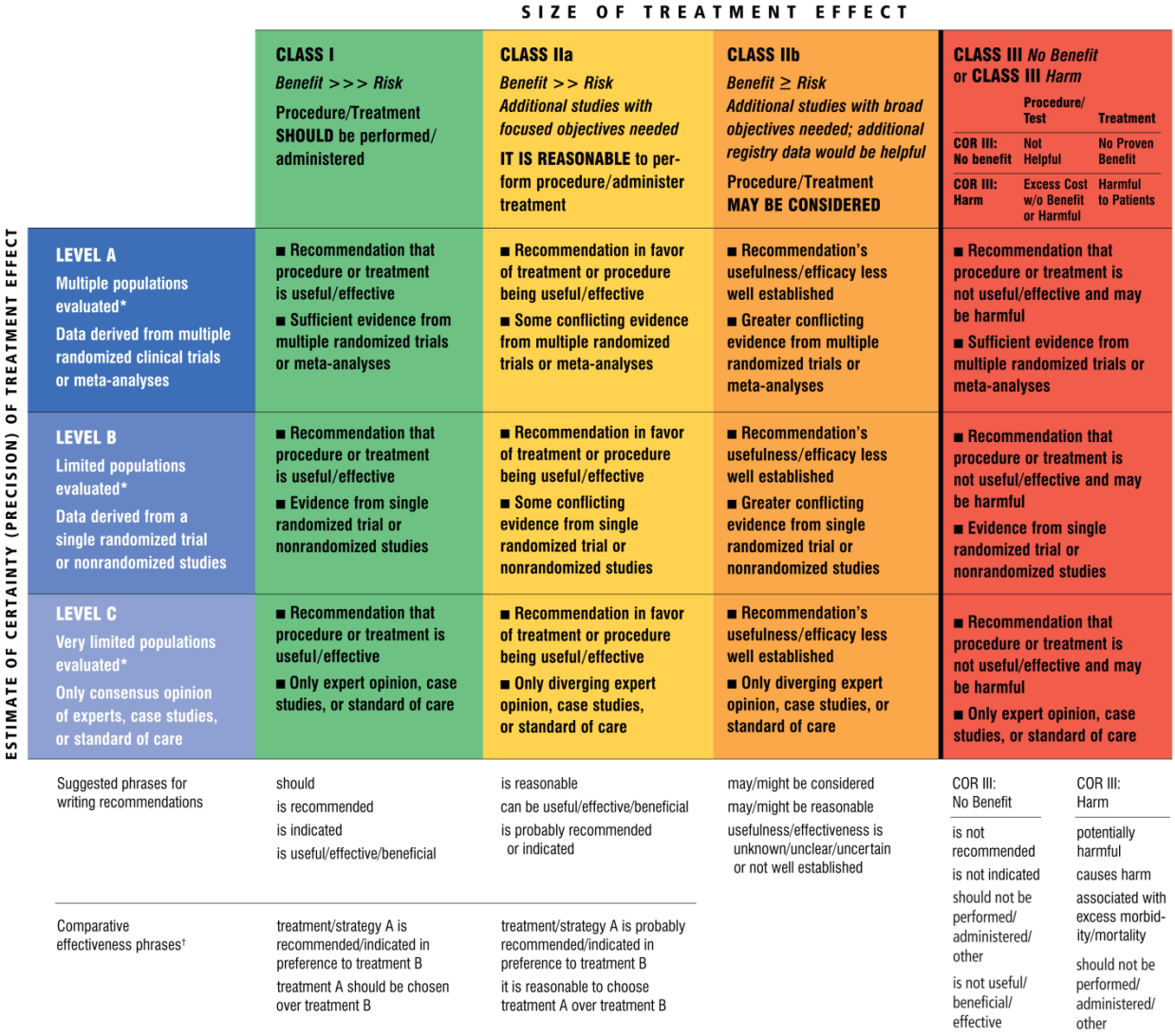
Class IIb: Procedure/Treatment may be considered

Class III: No benefit (Not helpful or No proven benefit)

Class III: Harm (Excess cost w/o benefit or Harmful to patients)

Specific COR definitions are included in Table 1 below.

Table 1. Applying Classification of Recommendation and Level of Evidence



Note: A recommendation with Level of Evidence B or C does not imply that the recommendation is weak. Many important clinical questions addressed in the guidelines do not lend themselves to clinical trials. Although randomized trials are unavailable, there may be a very clear clinical consensus that a particular test or therapy is useful or effective. *Data available from clinical trials or registries about the usefulness/efficacy in different subpopulations, such as sex, age, history of diabetes, history of prior myocardial infarction, history of heart failure, and prior aspirin use. †For comparative effectiveness recommendations (Class I and IIa; Level of Evidence A and B only), studies that support the use of comparator verbs should involve direct comparisons of the treatments or strategies being evaluated.

1a.4.5. Citation and URL for methodology for grading recommendations (*if different from 1a.4.1*)**:**

ACCF/AHA Task Force on Practice Guidelines. Methodology Manual and Policies From the ACCF/AHA Task Force on Practice Guidelines. American College of Cardiology Foundation and American Heart Association, Inc. Cardiosource.com. 2010. Available at: http://assets.cardiosource.com/Methodology_Manual_for_ACC_AHA_Writing_Committees.pdf and http://my.americanheart.org/idc/groups/ahamah-public/@wcm/@sop/documents/downloadable/ucm_319826.pdf

1a.4.6. If guideline is evidence-based (rather than expert opinion), are the details of the quantity, quality, and consistency of the body of evidence available (e.g., evidence tables)?

☒ Yes → ***complete section 1a.7***

☐ No → ***report on another systematic review of the evidence in sections 1a.6 and 1a.7; if another review does not exist, provide what is known from the guideline review of evidence in 1a.7***

1a.5. UNITED STATES PREVENTIVE SERVICES TASK FORCE RECOMMENDATION

1a.5.1. Recommendation citation (*including date*) and **URL for recommendation** (*if available online*):

1a.5.2. Identify recommendation number and/or page number and **quote verbatim, the specific recommendation**.

1a.5.3. Grade assigned to the quoted recommendation with definition of the grade:

1a.5.4. Provide all other grades and associated definitions for recommendations in the grading system. (*Note: the grading system for the evidence should be reported in section 1a.7.*)

1a.5.5. Citation and URL for methodology for grading recommendations (*if different from 1a.5.1*)**:**

Complete section 1a.7

1a.6. OTHER SYSTEMATIC REVIEW OF THE BODY OF EVIDENCE

1a.6.1. Citation (*including date*) and **URL** (*if available online*):

1a.6.2. Citation and URL for methodology for evidence review and grading (*if different from 1a.6.1*)**:**

Complete section 1a.7

1a.7. FINDINGS FROM SYSTEMATIC REVIEW OF BODY OF THE EVIDENCE SUPPORTING THE MEASURE

If more than one systematic review of the evidence is identified above, you may choose to summarize the one (or more) for which the best information is available to provide a summary of the quantity, quality, and consistency of the body of evidence. Be sure to identify which review is the basis of the responses in this section and if more than one, provide a separate response for each review.

1a.7.1. What was the specific structure, treatment, intervention, service, or intermediate outcome addressed in the evidence review?

2013 ACCF/AHA Guideline for the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults (e2893-2894)

This guideline specifically reviewed the underlying evidence for the treatment of blood cholesterol levels to reduce ASCVD risk.

The members of the Expert Panel acknowledge the important contributions arising from decades of genetic and biochemical studies, observational epidemiological and ecological studies, and in vitro and animal experiments that associated higher low-density lipoprotein cholesterol (LDL-C) levels with greater ASCVD risk. These studies provided the rationale for RCTs, which in turn demonstrated that lowering cholesterol levels reduced ASCVD events and thereby established a central, causal role of atherogenic cholesterol-containing lipoprotein particles, particularly LDL, in the genesis of CHD and ASCVD.

Other strategies for using drug therapy to reduce ASCVD events have been advocated, including treat- to-cholesterol target, lowest-is-best, and risk-based treatment approaches. However, only 1 approach has been evaluated in multiple RCTs – the use of fixed doses of cholesterol-lowering drugs to reduce ASCVD risk. Because the overwhelming body of evidence came from statin RCTs, the Expert Panel appropriately focused on these statin RCTs to develop evidence-based guidelines for the reduction of ASCVD risk.

1a.7.2. Grade assigned for the quality of the quoted evidence with definition of the grade:

An overall grade for the quality of evidence was not assigned. Rather, the quality of a study (or set of studies) supporting a recommendation was graded on an estimate of the certainty or precision of the treatment effect (see 1a.4.3).

Recommendations used to support this measure have a:

- Level of Evidence of A: Data derived from multiple randomized clinical trials or meta-analyses. References used to determine level of evidence must be provided and cited with the recommendation.

OR

- Level of Evidence B: Data derived from a single randomized trial, or nonrandomized studies. References used to determine level of evidence must be provided and cited with the recommendation.

1a.7.3. Provide all other grades and associated definitions for strength of the evidence in the grading system.

Level of Evidence (LOE) is an estimate of the certainty or precision of the treatment effect.

Level of Evidence A: Data derived from multiple randomized clinical trials or meta-analyses. References used to determine level of evidence must be provided and cited with the recommendation.
Level of Evidence B: Data derived from a single randomized trial, or nonrandomized studies. References used to determine level of evidence must be provided and cited with the recommendation.
Level of Evidence C: Consensus opinion of experts, case studies, or standard of care.

Specific LOE definitions are included in Table 1 in 1a.4.4.

1a.7.4. What is the time period covered by the body of evidence? (*provide the date range, e.g., 1990-2010*). Date range: 1975-2009 However, RCTs with hard ASCVD outcomes of myocardial infarction (MI), stroke, and cardiovascular death published after that date range were eligible for consideration until the Expert Panel began deliberations on relevant recommendations (July 2013).

QUANTITY AND QUALITY OF BODY OF EVIDENCE

1a.7.5. How many and what type of study designs are included in the body of evidence? (*e.g., 3 randomized controlled trials and 1 observational study*)

The systematic review was limited to RCTs with ASCVD outcomes and systematic reviews and meta-analyses of RCTs with ASCVD outcomes. The body of evidence supporting the recommendations on ASCVD risk reduction in the secondary prevention setting includes 19 secondary prevention randomized controlled trials and 4 meta-analyses (e38-39).

1a.7.6. What is the overall quality of evidence across studies in the body of evidence? (*discuss the certainty or confidence in the estimates of effect particularly in relation to study factors such as design flaws, imprecision due to small numbers, indirectness of studies to the measure focus or target population*)

e39-43: **Recommendation 1** is based on high-quality evidence from three secondary prevention RCTs that support the use of high-intensity statin therapy in individuals with clinical ASCVD. These RCTs, which were conducted in individuals with CHD, showed that atorvastatin 80 mg reduced ASCVD events more than moderate-intensity statin therapy did (see evidence statement 6). Atorvastatin 80 mg, compared with placebo, also reduced ASCVD events in individuals with a history of stroke or TIA (see evidence statement 7). No titration to a specific LDL–C goal occurred in these trials (see evidence statement 1). High-intensity statin therapy was similarly efficacious in reducing ASCVD events in women and in men with established ASCVD (see evidence statement 12).

A high level of evidence from the 2010 CTT meta-analysis of 26 statin RCTs showed that the reduction in cardiovascular events was proportional to the average magnitude of LDL–C reduction; that is, cardiovascular event rates decreased by approximately 20% for each 39 mg/dL (1 mmol/L) reduction in LDL–C (see evidence statements 14, 19, and 21 to 23). A moderate level of evidence showed no other systematic difference among the trials after data were adjusted for the degree by which LDL–C was lowered (see evidence statement 26). On the basis of these data, a moderately high level of evidence suggests that the reduction in ASCVD from statin therapy is a class effect related to the magnitude of LDL–C reduction.

Therefore, moderate evidence supports the use of statins, other than atorvastatin 80 mg, that lower LDL–C by a magnitude similar to that seen with atorvastatin 40 or 80 mg. On average, atorvastatin 80 mg lowers LDL–C by at least

50%, compared with placebo (see evidence statement 7). However, in individuals who are unable to tolerate atorvastatin 80 mg, other dosages or statins that lower LDL–C by approximately 50%, such as atorvastatin 40 mg or rosuvastatin 20 mg, can be used. Rosuvastatin 20 mg reduced ASCVD risk in a primary- prevention population (see evidence statement 35), but it has not yet been studied in ASCVD outcome trials in secondary-prevention populations. Nonetheless, because a high level of evidence that the relative reduction in ASCVD risk is related to the magnitude of LDL–C reduction in individuals with CHD, acute coronary syndromes, or other CVD, in primary prevention settings, and in various patient subgroups (see evidence statements 8, 10 to 20, and 28 to 30) the Panel concludes that a high level of evidence supports the generalization of the efficacy demonstrated in one prevention setting to other prevention settings. No ASCVD outcomes trials using rosuvastatin 40 mg, the highest FDA-approved dose of rosuvastatin, were identified in the systematic review.

Because the three trials of atorvastatin 80 mg excluded individuals older than 75 or 80 years, or included few individuals older than 75, there are few data regarding the efficacy and safety of high-intensity statin therapy for individuals in this older age group. In the five trials comparing more more-intensive versus less less-intensive statin therapy in the CTT meta-analysis in participants older than 75, CVD risk reduction per 39 mg/dL (1 mmol/L) reduction in LDL–C was not significant, although there was no evidence of heterogeneity among these participants compared with participants younger than 65 and those aged 65 to 74 (see evidence statements 13). The decision to initiate statin therapy for CVD prevention in patients older than 75 is therefore based on extrapolation from clinical trial evidence and should be made on an individual basis (see Special Populations). Safety also might be a consideration because of an increasing number of comorbidities in older persons (see Safety). However, because the 75-year age limit in clinical trials represents age at entry, a high level of evidence supports continuation of statins beyond age 75 in persons who are already tolerating these drugs (see evidence statement 13). (Class I; **Level of Evidence A**)

Recommendation 2. A high level of evidence supports the use of moderate-intensity statin therapy for the secondary prevention of ASCVD (see evidence statements 13 to 18, 20 to 24, 27, and 28). These statin doses reduced LDL–C by 25 to <50% (**Table 2**). Moderate-intensity statin therapy is therefore appropriate in individuals unable to tolerate high-intensity statin therapy, or when high-intensity statins are contraindicated. Simvastatin 40 mg, pravastatin 40 mg, and fluvastatin 40 mg twice daily reduced ASCVD events, compared with placebo, in secondary-prevention populations (see evidence statement 13).

A high level of evidence showed that similar relative risk reductions (RRRs) from statin therapy occurred for various subgroups of patients with ASCVD (see evidence statements 16 to 18, 20, and 29). In the 2010 CTT meta-analysis of 26 randomized trials, a moderate level of evidence indicated that similar RRRs occurred regardless of LDL–C level (see evidence statement 19) or other risk factors such as hypertension, blood pressure, body mass index, HDL–C or triglyceride level, smoking status, or glomerular filtration rate (see evidence 18 and 20). Unlike the more-intensive versus less-intensive RCTs, statin-versus-control RCTs (most of which evaluated moderate-intensity statins) clearly demonstrated a similar magnitude of RRR in CVD risk per 39 mg/dL reduction in individuals aged >75 years (see evidence statement 13). Therefore the Panel did not include an upper age limit for initiation of moderate- intensity statin therapy in individuals with ASCVD. (Class I; **Level of Evidence A**)

Recommendation 3. The Panel recommends that the decision to initiate or continue statin therapy in individuals aged >75 years be based on clinical judgment weighing benefits in reducing ASCVD events, harms, and patient preferences. However, as noted in Recommendation 1, it is reasonable to continue statin therapy after age 75 in patients who are already tolerating it. In the open-label Incremental Decrease in End Points Through Aggressive Lipid Lowering (IDEAL) trial, individuals aged 65 to 80 years at baseline reported more adverse events, and discontinuation rates were slightly higher in the older group receiving atorvastatin 80 mg (11.8%) than in individuals aged <65 years (7.9%), whereas discontinuation rates were similar between the two groups at the simvastatin 20 to 40 mg dose (4.1% among individuals aged 65 to 80 years versus 4.2% among individuals aged <65 years).

Clinicians may wish to consider additional factors, such as increasing comorbidities and changing clinical priorities with advancing age, in making treatment decisions in patients in this age group. See Special Populations, Individuals Aged >75 Years, for further discussion. (Class IIa; **Level of Evidence B**)

ESTIMATES OF BENEFIT AND CONSISTENCY ACROSS STUDIES IN BODY OF EVIDENCE

1a.7.7. What are the estimates of benefit—magnitude and direction of effect on outcome(s) across studies in the body of evidence? (*e.g., ranges of percentages or odds ratios for improvement/ decline across studies, results of meta-analysis, and statistical significance*)

p. e18: Potential ASCVD risk-reduction benefits. The absolute reduction in ASCVD events from moderate- or high-intensity statin therapy can be approximated by multiplying the estimated 10-year ASCVD risk by the anticipated relative-risk reduction from the intensity of statin initiated (~30% for moderate-intensity statin or ~45% for high-intensity statin therapy). The net ASCVD risk-reduction benefit is estimated from the number of potential ASCVD events prevented with a statin, compared to the number of potential excess adverse effects.

1a.7.8. What harms were studied and how do they affect the net benefit (benefits over harms)?

p. e18: Potential adverse effects. The excess risk of diabetes is the main consideration in ~0.1 excess cases per 100 individuals treated with a moderate-intensity statin for 1 year and ~0.3 excess cases per 100 individuals treated with a high-intensity statin for 1 year. In RCTs, both statin-treated and placebo-treated participants experienced the same rate of muscle symptoms. The actual rate of statin-related muscle symptoms in the clinical population is unclear. Muscle symptoms attributed to statin therapy should be evaluated.

UPDATE TO THE SYSTEMATIC REVIEW(S) OF THE BODY OF EVIDENCE

1a.7.9. If new studies have been conducted since the systematic review of the body of evidence, provide for each new study: 1) citation, 2) description, 3) results, 4) impact on conclusions of systematic review.

No new studies relevant to the focus of the ACC/AHA guideline were identified (as of December 29, 2015).

1a.8 OTHER SOURCE OF EVIDENCE

If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, please describe the evidence on which you are basing the performance measure.

1a.8.1 What process was used to identify the evidence?

1a.8.2. Provide the citation and summary for each piece of evidence.

1. Evidence, Performance Gap, Priority – Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. ***Measures must be judged to meet all subcriteria to pass this criterion and be evaluated against the remaining criteria.***

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form
ASCVDStatinEvidenceForm_04272016.docx

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- disparities in care across population groups.

1b.1. **Briefly explain the rationale for this measure** *(e.g., the benefits or improvements in quality envisioned by use of this measure)*
Use of statin therapy as secondary prevention for patients with clinical atherosclerotic cardiovascular disease can substantially reduce the risk of future cardiovascular events by approximately 30% for patients treated with moderate-intensity and 45% with high-intensity statins (Stone, 2013). Previously, performance measures have focused on managing low-density lipoprotein cholesterol (LDL-C) levels to reduce future atherosclerotic cardiovascular disease (ASCVD) risk. The 2013 American College of Cardiology and American Heart Association guideline release on the treatment of blood cholesterol to reduce ASCVD risk focused its

recommendations on statin use and not LDL-C targets due on the overwhelming body of randomized controlled trial evidence showing a clear link between the use of fixed doses of cholesterol-lowering drugs and ASCVD risk reduction. Specifically, the Cholesterol Treatment Trialists provided a comprehensive assessment of the benefits observed with statins. They undertook meta-analyses of individual participant data from 26 RCTs and demonstrated reduction in all-cause mortality, which was largely attributable to significant reductions in deaths due to CAD and other cardiac causes. The majority of studies in the aforementioned report included patients with known ASCVD. Of the 26 RCTs included, 5 trials (39,612 subjects, all of whom had CAD) compared more versus less intensive statin regimens. The trials demonstrated that more intensive regimens produced a highly significant 15% further reduction in major vascular events, driven by reductions in coronary death or nonfatal MI, coronary revascularization, and ischemic stroke. The investigators also found no significant effects observed on deaths due to cancer or other nonvascular causes or on cancer incidence, even at low LDL-C concentrations (Baigent, 2010; Stone, 2013). As such, this measure focuses on the broad clinical ASCVD population and whether optimal treatment (ideally high-intensity statins or moderate-intensity if a contraindication to high-intensity statins exists) is achieved.

References:

1. Baigent C, Blackwell L, Emberson J, et al. Efficacy and safety of more intensive lowering of LDL cholesterol: a meta-analysis of data from 170,000 participants in 26 randomised trials. Lancet. 2010;376: 1670–1681.

2. Stone NJ, Robinson JG, Lichtenstein AH, Bairey Merz CN, Blum CB, Eckel RH, Goldberg AC, Gordon D, Levy D, Lloyd-Jones DM, McBride P, Schwartz JS, Shero

3. ST, Smith SC Jr, Watson K, Wilson PWF. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2014;63:2889–2934.

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for endorsement maintenance. Include mean, std dev, min, max, interquartile range, scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included). This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*

This measure was developed and approved by the ACC and AHA in 2014.

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement.

In light of the current measure's focus on ensuring maximal statin therapy, some data exists that demonstrates that patients are not yet receiving optimal statin doses:

- In a secondary analysis of the Translational Research Investigating Underlying Disparities in Acute Myocardial Infarction Patients' Health Status (TRIUMPH) study, only 23% of 4,271 patients discharged alive following an acute MI were on maximal statin therapy, with substantial variability across hospitals (Arnold, 2014).
- A study in the Get With The Guidelines Registry of 65,396 patients with acute coronary syndromes (ACS) who were discharged with lipid-lowering agents found that only 38.3% were discharged with intensive lipid-lowering therapy (Javed, 2011).
- Jeevanantham and colleagues examined compliance of prescribing practices with the 2013 ACC/AHA blood cholesterol guidelines. While 100% of CAD patients and 76% of PAD patients were receiving a statin, titration to the high-intensity doses varied by diagnoses and age. Specifically, just over 38% of CAD patients aged 75 years and younger and approximately 25% of those CAD patients over the age of 75 years received high-intensity statins. For patients with PAD, the rate of those receiving high-intensity statins was only 13% for both age groups (Jeevanantham, 2015).

There are several publications that show statin therapy continues to be underutilized in the broader ASCVD population (Stone, 2013):

- A study from the REACH (Reduction of Atherothrombosis for Continued Health) Registry found that only 83% of ambulatory patients with known ASCVD were receiving lipid-lowering agents. (Kumar, 2009).
- A prospective study by Rabus and colleagues of 73 patients with angiographically diagnosed CAD found that only 44% received prescriptions for statins (Rabus,2008).
- Reports from the National Cardiovascular Data Registry PINNACLE Registry of ambulatory patients with CAD revealed that:
  - Only 66.5% (103,830 of 156,145) were receiving optimal medical therapy (OMT), including statins (Maddox, 2014)
  - 77.8% (30,160 of 38,775) were prescribed statins (Arnold, 2011)
  - Uninsured patients were 6% less likely to receive lipid-lowering therapy (Smolderen, 2013).
- Shah and colleagues reported that, among 292 patients from Olmstead County, MN, with incident acute myocardial infarction (MI), only 44% were still taking statins 3 years after their infarction (Shah, 2009).

• A study by Borden and colleagues involving patients in the National Cardiovascular Data Registry CathPCI Registry failed to show any significant improvement in the prescription of OMT after PCI following publication of the results of the Clinical Outcomes Utilizing Revascularization and Aggressive Drug Evaluation (COURAGE) study. Among all 467,211 patients (173,416 before [37.1%] and 293,795 after [62.9%] the COURAGE trial) who met the study criteria, the use of OMT at discharge following PCI before and after the COURAGE trial was 63.5% (95% confidence interval, 63.3% to 63.7%) and 66.0% (95% confidence interval, 65.8% to 66.1%), respectively ($p < 0.001$) (Borden, 2011).

• Data from the National Health and Nutrition Examination Survey (NHANES) found that approximately 70% of patients with peripheral artery disease (PAD) reported that they did not take a statin (Pande, 2011).

• Makowsky and colleagues found that for Canadian patients with a new diagnosis of PAD and no other cardiovascular disease only 44% were currently taking statins (Makowsky, 2011).

• A systematic literature review form 1999 to 2008 found that approximately 45% of patients with PAD were receiving a lipid-lowering agent (Flu, 2010).

• Gamboa and colleagues conducted a review of statin use in high-risk cardiovascular groups in the Reasons for Geographic and Racial Differences in Stroke (REGARDS) study and found that overall statin use was just over 58% in patients with coronary heart disease and 41.7% in those patients with a history stroke or abdominal aortic aneurysm (Gamboa, 2014).

References:

1. Arnold SV, Kosiborod M, Tang F, et al. Patterns of statin initiation, intensification, and maximization among patients hospitalized with an acute myocardial infarction. Circulation. 2014;129:1303–1309.

2. Arnold SV, Spertus JA, Tang F, et al. Statin use in outpatients with obstructive coronary artery disease. Circulation. 2011;124:2405-2410.

3. Borden WB, Redberg RF, Mushlin AI, et al. Patterns and intensity of medical therapy in patients undergoing percutaneous coronary intervention. JAMA. 2011;305: 1882–1889.

4. Flu HC, Tamsma JT, Lindeman JHN, Hamming JF, Lardenoye JHP. A systematic review of implementation of established recommended secondary prevention measures in patients with PAOD. Eur J Vasc Endovasc Surg. 2010;39:70-86.

5. Gamboa CM, Saffor MM, Levitan EB, et al. Statin under-use and low prevalence of LDL-C control among U.S. adults at high risk of coronary heart disease. Am J Med Sci. 2014;348:108-114.

6. Javed U, Deedwania PC, Bhatt DL, et al. Use of intensive lipid-lowering therapy in patients hospitalized with acute coronary syndrome: an analysis of 65,396 hospitalizations from 344 hospital participating in Get With The Guidelines (GWTG). Am Heart J. 2011; 161:418–424.

7. Jeevanantham V, Sharma S, Thapa R, et al. Disparities in lipid management in patients with peripheral artery disease versus coronary artery disease: comparison between ATP III and 2013 AHA/ACC guidelines. J Am Coll Cardiol. 2015;65:doi:10.1016/S0735-1097(15)61433-4.

8. Kumar A, Fonarow GC, Eagle KA, et al. Regional and practice variation in adherence to guideline recommendations for secondary and primary prevention among outpatients with atherothrombosis or risk factors in the United States: a report from the REACH Registry. Crit Pathw Cardiol. 2009;8:104-111.

9. Maddox TM, Chan PS, Spertus JA, et al. Variations in coronary artery disease secondary prevention prescriptions among outpatient cardiology practices: insights from the NCDR (National Cardiovascular Data Registry). J Am Coll Cardiol. 2014;63:539-546.

10. Makowsky M, McMurtry MS, Elton T, et al. Prevalence and treatment patterns of lower extremity peripheral arterial disease among patients at risk in ambulatory health settings. Can J Cardiol. 2011;27:389.e11-389.e18.

11. Pande RL, Perlstein TS, Beckman JA, Creager MA. Secondary prevention and mortality in peripheral artery disease. National health and nutrition examination study, 1999 to 2004. Circulation;2011;124:17-23.

12. Rabus SA, Izzettin FV, Sancar M, et al. Five-year follow-up of drug utilization for secondary prevention in coronary artery disease. Pharm World Sci. 2008;30:753-8.

13. Shah ND, Dunlay SM, Ting HH, et al. Long-term medication adherence after myocardial infarction: experience of a community. Am J Med. 2009;122. 961–13.

14. Smolderen KG, Spertus JA, Tang F, et al. Treatment differences by health insurance among outpatients with coronary artery disease: insights from the National Cardiovascular Data Registry. J Am Coll Cardiol. 2013;61:1069-75.

15. Stone NJ, Robinson JG, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force

on Practice Guidelines. J Am Coll Cardiol. 2014;63 (25 pt B):2889-2934.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for endorsement maintenance. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include.) This information also will be used to address the subcriterion on improvement (4b.1) under Usability and Use.*

In light of the current measure's focus on ensuring maximal statin therapy, some data exists that demonstrates that patients are not yet receiving optimal statin doses:

- In a secondary analysis of the Translational Research Investigating Underlying Disparities in Acute Myocardial Infarction Patients' Health Status (TRIUMPH) study, only 23% of 4,271 patients discharged alive following an acute MI were on maximal statin therapy, with substantial variability across hospitals (Arnold, 2014).
- A study in the Get With The Guidelines Registry of 65,396 patients with acute coronary syndromes (ACS) who were discharged with lipid-lowering agents found that only 38.3% were discharged with intensive lipid-lowering therapy (Javed, 2011).
- Jeevanantham and colleagues examined compliance of prescribing practices with the 2013 ACC/AHA blood cholesterol guidelines. While 100% of CAD patients and 76% of PAD patients were receiving a statin, titration to the high-intensity doses varied by diagnoses and age. Specifically, just over 38% of CAD patients aged 75 years and younger and approximately 25% of those CAD patients over the age of 75 years received high-intensity statins. For patients with PAD, the rate of those receiving high-intensity statins was only 13% for both age groups (Jeevanantham, 2015).

There are several publications that show statin therapy continues to be underutilized in the broader ASCVD population (Stone, 2013):

- A study from the REACH (Reduction of Atherothrombosis for Continued Health) Registry found that only 83% of ambulatory patients with known ASCVD were receiving lipid-lowering agents. (Kumar, 2009).
- A prospective study by Rabus and colleagues of 73 patients with angiographically diagnosed CAD found that only 44% received prescriptions for statins (Rabus,2008).
- Reports from the National Cardiovascular Data Registry PINNACLE Registry of ambulatory patients with CAD revealed that:
  - Only 66.5% (103,830 of 156,145) were receiving optimal medical therapy (OMT), including statins (Maddox, 2014)
  - 77.8% (30,160 of 38,775) were prescribed statins (Arnold, 2011)
  - Uninsured patients were 6% less likely to receive lipid-lowering therapy (Smolderen, 2013).
- Shah and colleagues reported that, among 292 patients from Olmstead County, MN, with incident acute myocardial infarction (MI), only 44% were still taking statins 3 years after their infarction (Shah, 2009).
- A study by Borden and colleagues involving patients in the National Cardiovascular Data Registry CathPCI Registry failed to show any significant improvement in the prescription of OMT after PCI following publication of the results of the Clinical Outcomes Utilizing Revascularization and Aggressive Drug Evaluation (COURAGE) study. Among all 467,211 patients (173,416 before [37.1%] and 293,795 after [62.9%] the COURAGE trial) who met the study criteria, the use of OMT at discharge following PCI before and after the COURAGE trial was 63.5% (95% confidence interval, 63.3% to 63.7%) and 66.0% (95% confidence interval, 65.8% to 66.1%), respectively ($p < 0.001$) (Borden, 2011).
- Data from the National Health and Nutrition Examination Survey (NHANES) found that approximately 70% of patients with peripheral artery disease (PAD) reported that they did not take a statin (Pande, 2011).
- Makowsky and colleagues found that for Canadian patients with a new diagnosis of PAD and no other cardiovascular disease only 44% were currently taking statins (Makowsky, 2011).
- A systematic literature review form 1999 to 2008 found that approximately 45% of patients with PAD were receiving a lipid-lowering agent (Flu, 2010).
- Gamboa and colleagues conducted a review of statin use in high-risk cardiovascular groups in the Reasons for Geographic and Racial Differences in Stroke (REGARDS) study and found that overall statin use was just over 58% in patients with coronary heart disease and 41.7% in those patients with a history stroke or abdominal aortic aneurysm (Gamboa, 2014).

References:

1. Arnold SV, Kosiborod M, Tang F, et al. Patterns of statin initiation, intensification, and maximization among patients hospitalized with an acute myocardial infarction. Circulation. 2014;129:1303–1309.

2. Arnold SV, Spertus JA, Tang F, et al. Statin use in outpatients with obstructive coronary artery disease. Circulation. 2011;124:2405-2410.

3. Borden WB, Redberg RF, Mushlin AI, et al. Patterns and intensity of medical therapy in patients undergoing percutaneous coronary intervention. JAMA. 2011;305: 1882–1889.

4. Flu HC, Tamsma JT, Lindeman JHN, Hamming JF, Lardenoye JHP. A systematic review of implementation of established

recommended secondary prevention measures in patients with PAOD. Eur J Vasc Endovasc Surg. 2010;39:70-86.

5. Gamboa CM, Saffor MM, Levitan EB, et al. Statin under-use and low prevalence of LDL-C control among U.S. adults at high risk of coronary heart disease. Am J Med Sci. 2014;348:108-114.

6. Javed U, Deedwania PC, Bhatt DL, et al. Use of intensive lipid-lowering therapy in patients hospitalized with acute coronary syndrome: an analysis of 65,396 hospitalizations from 344 hospital participating in Get With The Guidelines (GWTG). Am Heart J. 2011; 161:418–424.

7. Jeevanantham V, Sharma S, Thapa R, et al. Disparities in lipid management in patients with peripheral artery disease versus coronary artery disease: comparison between ATP III and 2013 AHA/ACC guidelines. J Am Coll Cardiol. 2015;65:doi:10.1016/S0735-1097(15)61433-4.

8. Kumar A, Fonarow GC, Eagle KA, et al. Regional and practice variation in adherence to guideline recommendations for secondary and primary prevention among outpatients with atherothrombosis or risk factors in the United States: a report from the REACH Registry. Crit Pathw Cardiol. 2009;8:104-111.

9. Maddox TM, Chan PS, Spertus JA, et al. Variations in coronary artery disease secondary prevention prescriptions among outpatient cardiology practices: insights from the NCDR (National Cardiovascular Data Registry). J Am Coll Cardiol. 2014;63:539-546.

10. Makowsky M, McMurtry MS, Elton T, et al. Prevalence and treatment patterns of lower extremity peripheral arterial disease among patients at risk in ambulatory health settings. Can J Cardiol. 2011;27:389.e11-389.e18.

11. Pande RL, Perlstein TS, Beckman JA, Creager MA. Secondary prevention and mortality in peripheral artery disease. National health and nutrition examination study, 1999 to 2004. Circulation;2011;124:17-23.

12. Rabus SA, Izzettin FV, Sancar M, et al. Five-year follow-up of drug utilization for secondary prevention in coronary artery disease. Pharm World Sci. 2008;30:753-8.

13. Shah ND, Dunlay SM, Ting HH, et al. Long-term medication adherence after myocardial infarction: experience of a community. Am J Med. 2009;122. 961–13.

14. Smolderen KG, Spertus JA, Tang F, et al. Treatment differences by health insurance among outpatients with coronary artery disease: insights from the National Cardiovascular Data Registry. J Am Coll Cardiol. 2013;61:1069-75.

15. Stone NJ, Robinson JG, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2014;63 (25 pt B):2889-2934.

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations.

There is limited published data on potential disparities in care for ASCVD patients receiving statin therapy:

- The Southern Community Cohort Study found that blacks were less likely than whites to report that they were receiving a statin (47% to 52%) (Lipworth, 2014).
- Qato and colleagues found similar results where blacks at high risk for cardiovascular disease reported statin use 38% of the time as opposed to 50% for whites (Qato, 2010).
- Differences in prescribing rates for statins were identified in medical record reviews across 23 states. Only 47% of African American men received a statin while close to 60% of Caucasian men were prescribed lipid-lowering agents (Massing, 2004).
- Uninsured patients were 6% less likely to receive lipid-lowering therapy (Smolderen, 2013).

References:

1. Lipworth L, Fazio S, Kabagambe EK, et al. A prospective study of statin use and mortality among 67,385 blacks and whites in the southeastern United States. Clinical Epidemiology. 2014;6:15-25.

2. Massing MW, Foley KA, Carter-Edwards L, Sueta CA, Alexander CM, Simpson RJ. Disparities in lipid management for African Americans and Caucasians with coronary artery disease: a national cross-sectional study. BMC Cardiovascular Disorders. 2004;4: doi:10.1186/1471-2261-4-15.

3. Qato DM, Lindau ST, Conti RM, Schumm LP, Alexandr GC. Racial and ethnic disparities in cardiovascular medication use among older adults in the United States. Pharmacoepidemiol Drug Saf. 2010;19:834-842.

4. Smolderen KG, Spertus JA, Tang F, et al. Treatment differences by health insurance among outpatients with coronary artery disease: insights from the National Cardiovascular Data Registry. J Am Coll Cardiol. 2013;61:1069-1075.

1c. High Priority (previously referred to as High Impact)

The measure addresses:

- a specific national health goal/priority identified by DHHS or the National Priorities Partnership convened by NQF; OR
- a demonstrated high-priority (high-impact) aspect of healthcare (e.g., affects large numbers of patients and/or has a substantial impact for a smaller population; leading cause of morbidity/mortality; high resource use (current and/or future); severity of illness; and severity of patient/societal consequences of poor quality).

1c.1. **Demonstrated high priority aspect of healthcare**

Affects large numbers, A leading cause of morbidity/mortality, Severity of illness

1c.2. **If Other:**

1c.3. **Provide epidemiologic or resource use data that demonstrates the measure addresses a high priority aspect of healthcare. List citations in 1c.4.**

The number of individuals who should be treated with statin therapy increased from over 43 million to 56 million based on the new ACC/AHA guidelines for the treatment of cholesterol and specifically those patients with existing cardiovascular disease would broaden to include an additional 2.4 million (Pencina, 2014). The new guideline determined that the use of statin therapy as secondary prevention for patients with clinical atherosclerotic cardiovascular disease can substantially reduce the risk of future cardiovascular events by approximately 30% for patients treated with moderate-intensity and 45% with high-intensity statins (Stone, 2013). In addition, estimates of the potential quality-adjusted life years (QALY) that can be achieved if the current 10-year ASCVD risk threshold (>7.5% risk threshold) outlined in the ACC/AHA guidelines show that the incremental cost-effectiveness ratio would be $37,000/QALY. These estimates would be even greater if the risk thresholds were lessened (Pandya, 2015).

1c.4. **Citations for data demonstrating high priority provided in 1a.3**

1. Pandya A, Sy S, Weinstein MC, Gaziano TA. Cost-effectiveness of 10-year risk thresholds for initiation of statin therapy for primary prevention of cardiovascular disease. JAMA. 2015;314:142-150.

2. Pencina MJ, Navar-Boggan AM, D'Agostino RB, et al. Application of new cholesterol guidelines to a population-based sample. NEngl J Med. 2014;370:1422-1431.

3. Stone NJ, Robinson JG, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol. 2014;63 (25 pt B):2889-2934.

1c.5. **If a PRO-PM (e.g. HRQoL/functional status, symptom/burden, experience with care, health-related behaviors), provide evidence that the target population values the measured PRO and finds it meaningful.** *(Describe how and from whom their input was obtained.)*

Not Applicable.

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the subcriteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. **Subject/Topic Area** *(check all the areas that apply):*

Cardiovascular, Cardiovascular : Acute Myocardial Infarction, Cardiovascular : Percutaneous Coronary Intervention (PCI), Neurology : Stroke/Transient Ischemic Attack (TIA)

De.6. **Cross Cutting Areas** *(check all the areas that apply):*

S.1. **Measure-specific Web Page** *(Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials. Do not enter a URL linking to a home page or to general information.)*
http://content.onlinejacc.org/article.aspx?articleid=2476226

S.2a. **If this is an eMeasure**, HQMF specifications must be attached. Attach the zipped output from the eMeasure authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications)
This is not an eMeasure **Attachment:**

S.2b. **Data Dictionary, Code Table, or Value Sets** *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*
No data dictionary **Attachment:**

S.3. **For endorsement maintenance**, please briefly describe any changes to the measure specifications since last endorsement date and explain the reasons.
Not Applicable

S.4. **Numerator Statement** *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome)*
IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
Patients in the denominator who have been offered* high-intensity statin† OR have been offered* moderate-intensity statin†.

Definitions:
*A statin is "offered" if it is prescribed or if a patient reason exception for not being prescribed a statin is documented.

†Moderate-intensity and high-intensity statin doses are defined in Table 5 of the 2013 ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adult
http://content.onlinejacc.org/article.aspx?articleid=1879710)

S.5. **Time Period for Data** *(What is the time period in which data will be aggregated for the measure, e.g., 12 mo, 3 years, look back to August for flu vaccination? Note if there are different time periods for the numerator and denominator.)*
Reporting Year

S.6. **Numerator Details** *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm.
See Supplemental Resources attached in Appendix Field A.1.

S.7. **Denominator Statement** *(Brief, narrative description of the target population being measured)*
All patients 18-75 years of age with clinical ASCVD* who were seen within a 12-month period. This measure is designed to apply to chronic care populations and does not apply to patients in acute care hospitals.

Definition:
*Clinical ASCVD includes acute coronary artery syndromes, a history of MI, stable or unstable angina, coronary or other arterial revascularization, stroke, TIA, and peripheral arterial disease presumed to be of atherosclerotic origin.

S.8. **Target Population Category** *(Check all the populations for which the measure is specified and tested if any):*
Populations at Risk

S.9. **Denominator Details** *(All information required to identify and calculate the target population/denominator such as definitions, specific data collection items/responses , code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
See 'Registry Supplemental Resources' attached in appendix field A.1 for data dictionary and form.

S.10. **Denominator Exclusions** *(Brief narrative description of exclusions from the target population)*
Exceptions: Documentation of medical reason(s) for not prescribing a statin (e.g., allergy, intolerance to statin[s], other medical reasons).

S.11. **Denominator Exclusion Details** *(All information required to identify and calculate exclusions from the denominator such as definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at S.2b)*
The ACC and AHA distinguish between measure exceptions and measure exclusions. Exclusions arise when the intervention required by the numerator is not appropriate for a group of patients who are otherwise included in the initial patient or eligible population of a measure (i.e. the denominator). Exclusions are absolute and are to be removed from the denominator of a measure and therefore clinical judgment does not enter the decision.

Measure Exceptions:

Denominator Exceptions are used to remove a patient from the denominator of the [performance measure when the patient does no receive a therapy or service AND that therapy or service would not be appropriate due to the patient specific reasons, the patient would otherwise meet the denominator. Exceptions are not absolute, and are based on clinical judgment, individual patient characteristics or patients. For this measure exception may include medical reasons for not prescribing a statin (e.g., allergy, intolerance to statin[s],other medical reasons). There are no patient or system reasons that would remove a patient from the denominator.

S.12. **Stratification Details/Variables** *(All information required to stratify the measure results including the stratification variables, definitions, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format with at S.2b)*
We encourage that the results of this measure be stratified by race, ethnicity, administrative sex, and payer, consistent with the data elements collected in the Pinnacle Registry.

S.13. **Risk Adjustment Type** (Select type. Provide specifications for risk stratification in S.12 and for statistical model in S.14-15)
No risk adjustment or risk stratification
If other:

S.14. **Identify the statistical risk model method and variables** *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development and testing should be addressed with measure testing under Scientific Acceptability)*
Not Applicable

S.15. **Detailed risk model specifications** *(must be in attached data dictionary/code list Excel or csv file. Also indicate if available at measure-specific URL identified in S.1.)*
Note: Risk model details (including coefficients, equations, codes with descriptors, definitions), should be provided on a separate worksheet in the suggested format in the Excel or csv file with data dictionary/code lists at S.2b.

S.15a. **Detailed risk model specifications** (*if not provided in excel or csv file at S.2b*)
Not Applicable.

S.16. **Type of score:**
Rate/proportion
If other:

S.17. **Interpretation of Score** *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*
Better quality = Higher score

S.18. **Calculation Algorithm/Measure Logic** *(Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.)*
To calculate performance rates:

1) Find the patients who meet the initial patient population (i.e., the general group of patients that a set of performance measures is designed to address).
2) From the patients within the initial patient population criteria, find the patients who qualify for the denominator. (i.e., the specific group of patients for inclusion in a specific performance measure based on defined criteria). Note: in some cases the initial patient population and denominator are identical.
3) Find the patients who quality for exclusions and subtract from the denominator.
4) From the patients within the denominator (after exclusions have been subtracted from the denominator), find the patients who qualify for the Numerator (i.e., the group of patients in the denominator for whom a process or outcome of care occurs). Validate that the number of patients in the numerator is less than or equal to the number of patients in the denominator.
5) From the patients who did not meet the numerator criteria, determine if the physician has documented that the patient meets any criteria for exception when exceptions have been specified [for this measure: medical reason(s) (e.g., allergy, intolerance to statin[s], other medical reasons)]. If the patient meets any exception criteria, they should be removed from the denominator for performance calculation. --Although the exception cases are removed from the denominator population for the performance calculation, the exception rate (i.e., percentage of patients with valid exceptions) should be calculated and reported along with performance rates to track variations in care and highlight possible areas of focus for QI.

If the patient does not meet the numerator and a valid exception is not present, this case represents a quality failure.

For calculation algorithm, see 'Registry Supplemental Resources' attached in appendix field A.1.

S.19. **Calculation Algorithm/Measure Logic Diagram URL or Attachment** *(You also may provide a diagram of the Calculation Algorithm/Measure Logic described above at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
Available in attached appendix at A.1

S.20. **Sampling** *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*
IF a PRO-PM, identify whether (and how) proxy responses are allowed.
Not Applicable, this measure is not based on a sample.

S.21. **Survey/Patient-reported data** *(If measure is based on a survey, provide instructions for conducting the survey and guidance on minimum response rate.)*
IF a PRO-PM, specify calculation of response rates to be reported with performance measure results.
Not applicable, this measure is not based on a survey.

S.22. **Missing data** (specify how missing data are handled, e.g., imputation, delete case.)
Required for Composites and PRO-PMs.
In PINNACLE missing values are interpreted as "no" for most variables. It is challenging to distinguish real missing versus "No." Therefore, we are assuming that missing documentation of statin therapy indicates a failure to meet the measure. It is possible that a provider may not have documented diagnoses of ASCVD use in their EMR system precluding inclusion of the patient in the denominator, although this seems unlikely. More importantly, there were some challenges in capturing statin dosing from some providers, as described in Section 1.2 of the testing form.

S.23. **Data Source** *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*
If other, please describe in S.24.
Electronic Clinical Data : Registry

S.24. **Data Source or Collection Instrument** *(Identify the specific data source/data collection instrument e.g. name of database, clinical registry, collection instrument, etc.)*
IF a PRO-PM, identify the specific PROM(s); and standard methods, modes, and languages of administration.
See 'Registry Supplemental Resources' attached in appendix field A.1.

S.25. **Data Source or Collection Instrument** *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*
Available in attached appendix at A.1

S.26. **Level of Analysis** *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*
Clinician : Individual

S.27. **Care Setting** *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*
Ambulatory Care : Clinician Office/Clinic
If other:

S.28. **COMPOSITE Performance Measure** - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*
Not Applicable

2a. Reliability – See attached Measure Testing Submission Form
2b. Validity – See attached Measure Testing Submission Form
ASCVDStatinTestingForm_4.27.16-635973658083446659.docx

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b2-2b7)

Measure Number (*if previously endorsed*)**:** Click here to enter NQF number
Measure Title: Statin Therapy in Patients with Clinical Atherosclerotic Disease
Date of Submission: 4/27/2016
Type of Measure:

☐ Composite – ***STOP – use composite testing form***	☐ Outcome (*including PRO-PM*)
☐ Cost/resource	☒ Process
☐ Efficiency	☐ Structure

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
- **For all measures, sections 1, 2a2, 2b2, 2b3, and 2b5 must be completed.**
- **For outcome and resource use measures**, section **2b4** also must be completed.
- If specified for **multiple data sources/sets of specificaitons** (e.g., claims and EHRs), section **2b6** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b2-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 20 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address sociodemographic variables and testing in this form refer to the release notes for version 6.6 of the Measure Testing Attachment.

Note: The information provided in this form is intended to aid the Steering Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a2. Reliability testing [10] demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **PRO-PMs and composite performance measures**, reliability should be demonstrated for the computed performance score.

2b2. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **PRO-PMs and composite performance measures**, validity should be demonstrated for the computed performance score.

2b3. Exclusions are supported by the clinical evidence; otherwise, they are supported by evidence of sufficient frequency of

occurrence so that results are distorted without the exclusion; [12]

AND

If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, 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). [13]

2b4. For outcome measures and other measures when indicated (e.g., resource use):

- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and sociodemographic factors) that influence the measured outcome and are present at start of care; [14,15] and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b5. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful [16] differences in performance**;

OR

there is evidence of overall less-than-optimal performance.

2b6. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b7. For **eMeasures, composites, and PRO-PMs** (or other measures susceptible to missing data), analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements 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 of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality.

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

13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions

15. 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 percent v. 75 percent) 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 less-than-optimal performance may not demonstrate much variability across providers.

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure*. *Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: **(*must be consistent with data sources entered in S.23*)**	**Measure Tested with Data From:**
☐ abstracted from paper record	☐ abstracted from paper record
☐ administrative claims	☐ administrative claims
☒ clinical database/registry	☒ clinical database/registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

1.2. If an existing dataset was used, identify the specific dataset (*the dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry*).

The primary analysis was conducted at the level of the individual provider and included all patients with atherosclerotic cardiovascular risk in adults with established atherosclerotic cardiovascular disease (ASCVD) cared for by that provider and captured in the PINNACLE Registry during the one-year study period. PINNACLE is an outpatient registry that maps data from the EHR of 206 cardiovascular practices. The PINNACLE Registry systematically maps each practice's Electronic Health Record to the data elements required for the PINNACLE Registry, with careful validation of the translation process prior to enrollment. Data from the registry are reported back to the practices on a quarterly basis for quality improvement. We would note that the Pinnacle Registry data harvest is an iterative and continuously improving process, and the assumption is that this will improve over time.

Using these data, we were able to calculate the number of patients who should have received statin therapy, or a clinically, evidence-based reason not to use statin therapy was documented. This means that every eligible patient in that provider's practice is included. For this measure, providers with less than 10 eligible patient encounters during the study period were excluded, since performance estimates are unstable with such small numbers. All other cases from all practices and providers were included. We included all visits for each patient and considered the performance measure met if they met the criterion for the measure at any encounter during the year.

This data extraction process is clearly feasible, but often requires iterative improvements to acquire all relevant data in all practices, particularly as the requisite data for new measures emerge. For example, in 2014, there were 2,503,852 patients treated by these 206 practices. However, we needed to exclude 68 practices (678,475 patients) because their EHR did not transmit the data on medication dose needed to determine whether the measure was met. In the future, this limitation will be addressed by remapping their data extraction. Within the remaining 138 practices, we excluded 841,376 patients in whom comorbidity data to determine the presences of ASCVD was missing. We also need to remove 205,345 patients from the remaining 138 practices for whom we knew they were on statins, but were unable to define the dose or they had a medical exclusion for statins. As opposed to the administrative limitation with the 68 practices that had no mapping of their statin dose data, this limitation (as does the preceding one) has the potential to introduce some selection bias into our assessment. Finally, we needed to remove an additional 16,188 patients and 2 practices because the providers treated <10 eligible patients and were thus not eligible for the measure, an exclusion that also should not introduce any selection biases. A similar rate of patient exclusions was also observed with the 2013 data, although we recognize that the performance in 2013 might be expected to be lower as the guideline was introduced during that calendar year. This left 239,948 patients from 136 practices for evaluation of the proposed performance

measure in 2014 and 209,770 patients from 121 practices in 2013. The evaluation of the measure's performance applies to these practices and patients and is described in more detail below.

1.3. What are the dates of the data used in testing? The primary analysis included encounters between 01/01/2014-12/31/2014. Additionally, we used data from 01/01/2013 thru 12/31/2013 for temporal comparison.

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (***must be consistent with levels entered in item S.26***)	**Measure Tested at Level of:**
☒ individual clinician	☒ individual clinician
☐ group/practice	☐ group/practice
☐ hospital/facility/agency	☐ hospital/facility/agency
☐ health plan	☐ health plan
☐ other: Click here to describe	☐ other: Click here to describe

1.5. How many and which measured entities were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample*)

2013

In 2013, 1,701 providers met the minimum number of eligible patients (10) for inclusion in the reliability analysis. The average number of eligible patients for providers included in the measurement is 123.2 for a total of 209,770 patients. The ratio of patients to providers ranged from 10 to 1,427.

The unit of analysis for this measure is the provider. A description of the providers studied for the 2013 calendar year is shown below:

	Total
	n = 1701
Provider gender	
Male	1379 (81.1%)
Female	321 (18.9%)
Missing (.)	1
Provider categories	
NP/PA	179 (11.1%)
MD/DO	1424 (88.0%)
RN/nurses	16 (1.0%)
Missing (.)	82
Region	
(1) Northeast	186 (10.9%)
(2) Midwest	347 (20.4%)
(3) South	796 (46.8%)
(4) West	372 (21.9%)

2014

In 2014, 1,890 providers met the minimum number of eligible patients (10) for inclusion in the reliability analysis. The average number of eligible patients for providers included in the assessment is 126.9 for a total of 239,948 patients. The ratio of patients to providers ranged from 10 to 879.

The unit of analysis for this measure is the provider. A description of the providers studied for the 2014 calendar year is shown below:

	Total n = 1890
Provider gender	
Male	1511 (80.0%)
Female	378 (20.0%)
Missing (.)	1
Provider categories	
NP/PA	206 (11.5%)
MD/DO	1563 (87.4%)
RN/nurses	20 (1.1%)
Missing (.)	101
Region	
(1) Northeast	205 (10.8%)
(2) Midwest	399 (21.1%)
(3) South	898 (47.5%)
(4) West	388 (20.5%)

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

2013

There are a total of 209,770 patients included in the temporal comparison that were treated in 2013. Patients' characteristics are provided below:

	Total n = 209770
Race (Derived)	
(1) White	120382 (88.7%)
(2) Black	12039 (8.9%)
(3) Other	3358 (2.5%)
Missing (.)	49488
Insurance (Derived)	
(0) No insurance	581 (0.7%)
(1) Private	69332 (80.7%)
(2) Medicare	14807 (17.2%)
(3) Medicaid	142 (0.2%)
(4) Other	1075 (1.3%)
Missing (.)	33750
Age	
18 to <60	64132 (30.6%)
60 to <70	84735 (40.4%)
70 to <80	60903 (29.0%)

Sex	
Male	124534 (59.4%)
Female	85206 (40.6%)
Missing (.)	30
BMI (kg/m2) (Derived)	36.4 ± 78.9
Missing	41974
Diabetes Mellitus (any)	75929 (40.9%)
Coronary Artery Disease	153870 (83.7%)
Hypertension	154207 (80.6%)
Atrial Fibrillation or Flutter	39802 (19.0%)
Heart Failure	51551 (25.5%)
Peripheral Arterial Disease (Unspecified)	52036 (26.6%)
Stroke/TIA (Derived)	18820 (11.0%)
Myocardial Infarction	60976 (31.8%)

2014

There are a total of 239,948 patients included in the temporal comparison that were treated in 2014. Patients' characteristics are provided below:

	Total **n = 239948**
Race (Derived)	
(1) White	138613 (89.2%)
(2) Black	13247 (8.5%)
(3) Other	3540 (2.3%)
Missing (.)	63562
Insurance (Derived)	
(0) No insurance	668 (0.8%)
(1) Private	71294 (82.1%)
(2) Medicare	13720 (15.8%)
(3) Medicaid	145 (0.2%)
(4) Other	1041 (1.2%)
Missing (.)	42883
Age	
18 to <60	68641 (28.6%)
60 to <70	98138 (40.9%)
70 to <80	73169 (30.5%)
Sex	
Male	143727 (59.9%)
Female	96201 (40.1%)
Missing (.)	20
BMI (kg/m2) (Derived)	36.9 ± 84.3
Missing	41968
Diabetes Mellitus (any)	90469 (41.9%)
Coronary Artery Disease	170076 (83.8%)
Hypertension	180915 (82.5%)
Atrial Fibrillation or Flutter	48835 (20.4%)

	Total
	n = 239948
Heart Failure	61083 (26.5%)
Peripheral Arterial Disease (Unspecified)	65383 (29.0%)
Stroke/TIA (Derived)	24428 (12.4%)
Myocardial Infarction	68060 (31.6%)

1.7. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing reported below.
The dataset described above was used for all aspects of testing.
1.8 What were the patient-level sociodemographic (SDS) variables that were available and analyzed in the data or sample used? For example, patient-reported data (e.g., income, education, language), proxy variables when SDS data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate).
We do not currently collect any of the SDS variables examples listed above. As is noted in other sections of this testing form we do collect data on race as well as insurance type.

_____

2a2. RELIABILITY TESTING
Note*: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)
☐ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)
☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level checked above, describe the method of reliability testing and what it tests (*describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used*)
Reliability of the computed measure score was measured as the ratio of signal to noise. The signal in this case is the proportion of the variability in measured performance that can be explained by real differences in physician performance. Reliability at the level of the specific physician is given by:
Reliability = Variance (physician-to-physician) / [Variance (physician-to-physician) + Variance (physician-specific-error].
Thus, the reliability is the ratio of the physician-to-physician variance divided by the sum of the physician-to-physician variance plus the error variance specific to a physician. A reliability of zero implies that all the variability in a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in physician performance.

Reliability testing was performed by using a beta-binomial model. The beta-binomial model assumes the physician performance score is a binomial random variable conditional on the physician's true value that comes from the beta distribution. The beta distribution is usually defined by two parameters, alpha and beta. Alpha and beta can be thought of as intermediate calculations to get to the needed variance estimates.

Reliability is estimated for five different points: at the minimum number of quality reporting events for the measure; at the mean number of quality reporting events per physician; and at the 25th, 50th and 75th percentiles of the number of quality reporting events.

2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)
2013- In 2013 the signal to-noise ratios are shown below:

Description	Number of Patients	Signal-to-Noise Ratio
Minimum	10	0.985
25th percentile	53	0.991
50th percentile	94	0.993
75th percentile	161	0.996
Average (mean)	124	0.995

2014- In 2014, the signal-to-noise ratios are shown below:

Description	Number of Patients	Signal-to-Noise Ratio
Minimum	10	0.986
25th percentile	57	0.992
50th percentile	102	0.994
75th percentile	168	0.996
Average (mean)	127	0.995

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i*.e., what do the results mean and what are the norms for the test conducted?*)
For this measure the reliability was very high and was similar for 2013 and 2014, supporting the reproducibility of these estimates across years. At the minimum number of patients visits required (>10) the average reliability was 0.985 and 0.986, for 2013 and 2014 respectively. For providers with the median number of patients encounters, the reliability was 0.993 (2013) and 0.994 (2014). Given that a reliability of 0.70 is generally considered a minimum threshold for acceptability, and 0.80 is considered very good reliability, these data suggest that the measure is exceedingly good at describing differences across physicians.

2b2. VALIDITY TESTING
2b2.1. What level of validity testing was conducted? (*may be one or both levels*)
☐ **Critical data elements** (*data element validity must address ALL critical data elements*)
☒ **Performance measure score**
☐ **Empirical validity testing**
☒ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*)
2b2.2. For each level of testing checked above, describe the method of validity testing and what it tests (*describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used)*

Content validity for this measure was assessed by an expert workgroup through an iterative development process and was based upon the 2013 ACC/AHA Cholesterol Guidelines. Unlike the preceding guideline, the Adult Treatment Panel for the Detection, Evaluation and Treatment of High Blood Cholesterol in Adults – 3rd revision (ATP –III), the 2013 ACC/AHA guideline found evidence mostly for moderate to high intensity statin therapy among all lipid lowering therapy and recommended a fixed moderate to high intensity statin therapy in patients with ASCVD. Whereas the ATP-III guideline recommended use of both statin and non-statin lipid lowering therapy to achieve LDL-C target, this was an extrapolation from clinical trials, which tested only a fixed statin dose strategy to lower cholesterol. No clinical trials had tested a strategy to achieve LDL-C target and therefore the 2013 ACC/AHA guideline abolished an LDL-C target. It was expected

that the simplification of cholesterol management with the new guideline would encourage providers to streamline use of statin therapy rather than trying to achieve a specific level of LDL-C. The latter approach could also encourage providers to use multiple lipid lowering therapies and not primarily use statin therapy, which had the highest evidence as compared with non-statin therapy that did not have evidence (when the guideline was released) in reducing ischemic cardiovascular events. Furthermore, the new guideline highlighted the importance of discussion about benefits and harms of statin therapy between patient and providers, emphasizing the role of shared decision making in cholesterol management.

Based upon this guideline, this performance measure was designed to be integrated across multiple performance measurement sets impacting patients with atherosclerotic coronary disease, including stable ischemic heart disease, acute myocardial infarction, percutaneous coronary interventions and peripheral artery disease. Workgroup members thus included the chairs of the ACC/AHA work groups that had developed performance measure sets for coronary artery disease (Joseph P. Drozda, Jr., MD, FACC), peripheral arterial disease (Jeffrey W. Olin, DO, FACC, FAHA), percutaneous coronary intervention (Brahmajee K. Nallamothu, MD, FACC), and acute myocardial infarction (Harlan M. Krumholz, MD, SM, FACC). Also on the workgroup were T. Bruce Ferguson, Jr., MD, FACC, FAHA (cardiac surgeon), Hani Jneid, MD, FACC, FAHA, FSCAI (interventional cardiologist), and Henry H. Ting, MD, MBA, FACC, FAHA (health services and quality of care researcher). The primary mission of the workgroup was to develop updates of the lipid management measures in the CAD, PAD, PCI, and AMI measures sets based on the ACC/AHA 2013 lipid performance measures. The workgroup sought to develop a measure that was based on the strongest guideline recommendations for patients with established ASCVD. The revised guidelines emphasize a targeted treatment approach (based on patients' risk and absolute risk reduction) rather than a treat-to-target approach of achieving a certain LDL target. This led to a focus both on the type of lipid lowering (statins vs. other lipid lowering regimens) and the strength (moderate- to high-intensity doses) of statin dosing. Additional input on the content validity of draft measures was established through a 30-day public comment period and concurrent formal peer review process. All comments received were reviewed by the expert work group and the measures were adjusted to reflect the provided input. Additionally, the measure underwent review and approval by the Board of Trustees of the ACC and the Science Advisory and Coordinating Committee of the AHA, as well as review by a PCPI representative (Richard Hellman, MD, FACP, FACE – endocrinologist).

Construct Validity was difficult to establish because there has not been an independent audit of these data. However, it is important to note that an independent audit would merely involve an abstractor reviewing the same medical record from which PINNACLE directly abstracts its data and, given the identical source of the data, any error observed would either be due to the auditor incorrectly abstracting the data from the EHR or PINNACLE incorrectly mapping the data elements from the EHR. To address the latter, we conduct detailed analyses to insure that this does not happen and quarantine (i.e. not report) data that fails our Data Quality Review process. Validity of measure data elements in PINNACLE is routinely evaluated on a quarterly basis as part of the standard data extraction and analytic data set creation process. First, all relevant data elements are reviewed at the record level to ensure that individual data values are valid; any invalid values are set to missing. Next, the *distribution* of each data element is reviewed, aggregating both across practices and across calendar quarters within each practice, to identify outliers, suspicious patterns and/or systematic changes in the prevalence of the data element that may suggest data mapping errors or unanticipated changes in definitions, coding consistency, data completeness, etc. Identification of suspicious data includes both statistical criteria, using quality control charts with rigorous definitions of "out of control" rates, and manual clinical review of each distribution for plausibility. Records that are flagged as suspicious by these criteria are quarantined and excluded from analysis and reporting. The patients and practices removed are detailed in section 1.2 above. Importantly, there are ongoing efforts to work with each practice to improve data mapping and we believe that this will improve markedly over time. Importantly, because most of these exclusions are due to mapping difficulties at the practice level, rather than for specific patients within each practice, we believe that these administrative errors do not introduce bias in those for whom the measure can be calculated.

Face validity of the measure score was systematically assessed as follows:

After the measure was fully specified, members of two existing committees, one at the ACC and one at AHA, with expertise in in general cardiology, interventional cardiology, heart failure, electrophysiology and quality improvement, outcomes research, informatics and performance measurement, who were not involved in development of the measure, were asked to review the measure specifications and rate their agreement with the following statement:

"The scores obtained from the measure as specified will provide an accurate reflection of quality and can be used to distinguish good and poor quality."

The respondents recorded their rating on a scale of 1-5, where 1= Strongly Disagree; 3=Neither Agree nor Disagree; 5= Strongly Agree

8 members completed the survey and provided a mean importance rating of 4.63, with 100% agreeing with the use of the measure for quality assessment.

2b2.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)
We believe that the processes used to extract data from the exact source from which any abstraction process done manually would use (i.e. the EHR), and our thorough data quality review, provide strong evidence for the validity of this measure.

The results of the expert panel rating of the validity statement were as follows:

N = 8; Mean rating = 4.63 and 100% of respondents either agree or strongly agree that this measure can accurately distinguish good and poor quality.

Frequency Distribution of Ratings
1 - 0 (Strongly Disagree)
2 - 0
3 - 0 (Neither Agree nor Disagree)
4 - 3
5 - 5 (Strongly Agree)

2b2.4. What is your interpretation of the results in terms of demonstrating validity? (i*.e., what do the results mean and what are the norms for the test conducted?*)
The measure was judged to have high face validity by both its clinical importance and by the independent group of experts asked to rate it. The majority of experts agreed that the measure, as specified, will provide an accurate reflection of quality and can be used to distinguish good and poor quality. Importantly, as a process measure, the strong association of treatment with improved survival and reduced myocardial infarction rates provide strong validity for this measure as a mechanism to insure that strong clinical evidence is being translated to routine clinical care.

2b3. EXCLUSIONS ANALYSIS
NA ☐ no exclusions — *skip to section 2b4*

2b3.1. Describe the method of testing exclusions and what it tests (*describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used*)
Since not all patients with atherosclerotic cardiovascular risk meet the guideline recommendations for statin therapy, exclusions in this measure are intended to remove patients from whom statin therapy may not be appropriate. For this measure, the only exclusion was a documented contraindication or allergy to simvastatin, atorvastatin or rosuvastatin. No other exclusions, such as patient preference or system reasons for not being treated were applied.
2b3.2. What were the statistical results from testing exclusions? (*include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores*)
For this measure, the only exclusion was for a documented medical contraindication for statin therapy. No patients, in either 2013 or 2014, had such a contraindication. Although there were a few patients identified as having patient-centered reasons for not receiving statins, these were not considered as exclusions from this performance measure.
2b3.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., the value outweighs the burden of increased data collection and analysis. Note: **If patient preference is an exclusion**, the measure must be specified so that the effect on the performance score is*

transparent, e.g., scores with and without exclusion)
We do not view any concerns with the exceptions for this measure. We do wish to raise one potential opportunity for misclassification of patients meeting the numerator of this measure. Some patients are unable to tolerate higher doses of stain therapy and are only treated with low-dose statins. These patients would not necessarily have a patient-level exclusion documented for not being treated with higher dose statins and would be misclassified as not meeting the numerator, even though higher dose statins were attempted by the clinician. Overcoming this limitation will require new fields be entered into EHRs to document these reasons, but this would not be well received by many practitioners who are already overwhelmed with EHR-required documentation. The consequence of such misclassification would be a slightly lower performance rate among some physicians. However, the major threat to this is if there is a preferential risk of some physicians treating more patients with intolerance to high-dose statins than other physicians. This does not seem likely. Moreover, we suspect that some physicians work harder to uptitrate statins than other physicians and should be credited for these efforts. The current process of calculating this performance measure provides those rewards to these physicians and enables patients to be treated with the doses of statins that confer the greatest clinical benefits. Finally, the proportion of patients tolerant of lower-dose, but not higher-dose, statins is certainly much smaller than the large number of patients currently not being treated with higher-dose statins and it is very unlikely that one practitioner having more intolerant patients than another would influence all physicians opportunity to improve.

_____

2b4. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES
If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b4.1. What method of controlling for differences in case mix is used?
☒ **No risk adjustment or stratification**
☐ **Statistical risk model with** Click here to enter number of factors **risk factors**
☐ **Stratification by** Click here to enter number of categories **risk categories**
☐ **Other,** Click here to enter description
2b4.2. If an outcome or resource use measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

2b4.3. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or sociodemographic factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*)

2b4.4a. What were the statistical results of the analyses used to select risk factors?
2b4.4b. Describe the analyses and interpretation resulting in the decision to select SDS factors (e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects)
2b4.5. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (*describe the steps—do not just name a method; what statistical analysis was used*)
Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.
If stratified, skip to 2b4.9
2b4.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**
2b4.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):
2b4.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:
2b4.9. Results of Risk Stratification Analysis:
2b4.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for the test conducted*)
2b4.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

2b5. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b5.1. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified (*describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided related to performance gap in 1b)*

We examined variation in provider performance on this measure based on sex, age, race, and a number of other patient factors to identify variations. The findings are represented for 2013 and 2014 respectively.

2013:

label	# of providers	# of patients	Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std Dev
Male	1701	124534	0.00%	4.05%	18.0%	28.2%	82.6%	24.1%	17.3%
Female	1701	85206	0.00%	1.96%	13.8%	20.0%	100%	18.0%	15.0%
Age: <60	1698	64132	0.00%	0.00%	12.5%	19.0%	85.7%	19.0%	15.0%
Age: 60 -< 70	1701	84735	0.00%	3.70%	17.9%	28.3%	85.7%	24.6%	17.4%
Age: 70 -< 75	1699	60903	0.00%	2.63%	18.0%	28.6%	88.9%	25.9%	18.0%
Insurance: None	187	581	0.00%	0.00%	4.29%	0.00%	100%	0.00%	17.6%
Insurance: Private	975	69332	0.00%	1.50%	17.4%	26.3%	100%	24.8%	17.9%
Insurance: Medicare	797	14807	0.00%	0.00%	6.09%	5.77%	100%	5.77%	14.5%
Insurance: Medicaid	32	142	0.00%	0.00%	8.26%	7.68%	100%	7.68%	20.2%
Insurance: Other	209	1075	0.00%	0.00%	9.15%	0.00%	100%	0.00%	25.3%
Race: White	1442	120382	0.00%	1.49%	14.6%	21.4%	88.9%	19.9%	16.1%
Race: Black	1177	12039	0.00%	0.00%	14.5%	22.2%	100%	22.2%	24.2%
Race: Other	749	3358	0.00%	0.00%	23.1%	44.4%	100%	44.4%	35.6%

2014:

label	# of providers	# of patients	Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std Dev
Overall	1890	239948	0.00%	8.33%	20.9%	29.0%	87.5%	20.7%	17.3%
Male	1889	143727	0.00%	8.77%	22.4%	31.9%	88.9%	23.2%	18.5%
Female	1890	96201	0.00%	6.02%	18.6%	25.9%	100%	19.9%	17.2%
Age: <60	1885	68641	0.00%	3.66%	17.0%	24.4%	100%	20.8%	17.9%
Age: 60 -< 70	1890	98138	0.00%	8.57%	22.5%	32.6%	92.3%	24.0%	18.7%
Age: 70 -< 75	1888	73169	0.00%	8.33%	22.3%	32.9%	100%	24.6%	18.8%
Insurance: None	133	668	0.00%	0.00%	3.84%	0.00%	100%	0.00%	15.9%
Insurance: Private	1020	71294	0.00%	10.0%	24.5%	33.9%	100%	23.9%	20.3%
Insurance: Medicare	788	13720	0.00%	0.00%	11.5%	14.3%	100%	14.3%	19.3%
Insurance: Medicaid	33	145	0.00%	0.00%	10.1%	10.0%	100%	10.0%	20.6%
Insurance: Other	198	1041	0.00%	0.00%	12.7%	5.26%	100%	5.26%	28.7%
Race: White	1617	138613	0.00%	6.41%	19.8%	28.6%	100%	22.2%	17.5%

label	# of providers	# of patients	Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std Dev
Race: Black	1312	13247	0.00%	0.00%	20.6%	33.3%	100%	33.3%	27.8%
Race: Other	881	3540	0.00%	0.00%	27.1%	50.0%	100%	50.0%	37.2%

These data stratify performance across important demographic and insurance groups as a foundation for evaluating current disparities in care. Examining the mean difference in the proportion of providers' patients meeting the performance measures, it is apparent that women, young patients (<60), uninsured and governmentally-insured (as compared with privately insured) patients and, to a lesser extent, white patients are less often treated according to current guideline recommendations. Holding providers accountable for this performance measure has a great opportunity to potentially eliminate these disparities in treatment. These data reflect similar patterns as seen in 2013, the year that the updated guideline was released.

2b5.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined*)

The data below for 2013 and 2014 respectively show an important variability in care and a substantial opportunity to improve the treatment of ASCVD patients to prevent subsequent cardiovascular events. For example, in 2014, we observed that among 1,890 providers, the average compliance with the guideline recommendations of moderate/high intensity statins in patients with established ASCVD was 20.9% with the lowest rate being 0% and a maximal rate of 87.5%. The top decile, an achievable benchmark of care, was 58.7%. While there may be some opportunity to improve performance by better documentation of medical reasons (such as statin intolerance), this is very unlikely to account for the magnitude of the opportunity to improve care. Moreover the variability across physicians is enormous and the use of a performance measure to improve the consistency of care represents an important opportunity to elevate the reliability of healthcare.

2013

Overall mean performance on this measure is 16.3%, with a standard deviation of 15.6%. The minimum score equals 0.00%, while the maximum score equals 81.4%. The interquartile range is 4.3% to 25% with the highest decile of performance being 50.1%.

1701 providers were measured, and the patient study sample equals 209,770. 59.4% of the sample is male. 88.7% of the sample is white, 8.9% is black, and 2.5 % identified as "other." The sample reached across all US regions, with 10.9% of providers in the Northeast, 20.4% of providers in the Midwest, 46.8% of providers in the South, and 21.9% of providers in the West.

# of providers	# of patients	Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std Dev
1701	209770	0.00%	4.30%	16.3%	25.0%	81.4%	20.7%	15.6%

	Mean
Decile 2	0.3%
Decile 3	4.1%
Decile 4	7.0%
Decile 5	10.2%
Decile 6	13.6%

	Mean
Decile 7	18.2%
Decile 8	24.9%
Decile 9	34.4%
Decile 10	50.1%

2014

Overall mean performance on this measure is 20.9%, with a standard deviation of 17.3%. The minimum score equals 0.00%, while the maximum score equals 87.5%. The interquartile range is 8.3% to 29.0% with the highest decile of performance being 58.7%.

1890 providers were measured, and the patient study sample equals 239,948. 59.9% of the sample is male. 89.2% of the sample is white, 8.5% is black, and 2.3 % identified as "other." The sample reached across all US regions, with 10.8% of providers in the Northeast, 21.1% of providers in the Midwest, 47.5% of providers in the South, and 20.5% of providers in the West.

# of providers	# of patients	Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std Dev
1890	239948	0.00%	8.33%	20.9%	29.0%	87.5%	20.7%	17.3%

	Mean
Decile 1	0.1%
Decile 2	4.2%
Decile 3	8.3%
Decile 4	11.5%
Decile 5	14.7%
Decile 6	18.4%
Decile 7	22.9%
Decile 8	29.1%
Decile 9	40.3%
Decile 10	58.7%

2b5.3. What is your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

2013: A large amount of variability was noted among providers. The performance-met rate range was 0-81.4% with the inter-quartile range being 4.3% to 25.0%. This yielded a Median Rate Ratio of 3.03 (2.89, 3.18). The Median Rate Ratio measures the variation between physician clusters by comparing 2 persons from two randomly chosen different physician clusters. A MRR of 3.03 indicates a greater than 3-fold average likelihood that a patient with ASCVD would be treated with moderate/high-dose statins by one provider as compared with another.

2014: A large amount of variability was also noted among providers in 2014. The performance-met rate range was 0-87.5% with the inter-quartile range being 8.3% to 29.0%. This yielded a Median Rate Ratio of 2.41(2.33, 2.50). A MRR of

2.41 indicates an almost 2.5-fold average likelihood (slightly smaller than in 2013) that a patient with ASCVD would be treated with moderate/high-dose statins by one provider as compared with another.

2b6. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS
***If only one set of specifications, this section can be skipped*.**

Note: *This item is directed to measures that are risk-adjusted (with or without SDS factors)* ***OR*** *to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eMeasures). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator).* ***Comparability is not required when comparing performance scores with and without SDS factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.***

2b6.1. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications (*describe the steps—do not just name a method; what statistical analysis was used*)

2b6.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

2b6.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

2b7. MISSING DATA ANALYSIS AND MINIMIZING BIAS
2b7.1. Describe the method of testing conducted to identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)
In PINNACLE missing values are interpreted as "no" for most variables. It is challenging to distinguish real missing versus "No." Therefore, we are assuming that missing documentation of statin therapy indicates a failure to meet the measure. It is possible that a provider may not have documented diagnoses of ASCVD use in their EMR system precluding inclusion of the patient in the denominator, although this seems unlikely. More importantly, there were some challenges in capturing statin dosing from some providers, as described in Section 1.2.
2b7.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)
Please see Section 1.2 for a description of the analytic cohort. By and large, most missing data was administrative and unlikely to introduce substantial bias in those in whom performance is reported. Among providers with statin dose obtained for some patients, but not others, the rate of missing data are shown below:
2013

Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std Dev
0.000	0.803	0.862	0.962	1.000	0.159	0.145

2014

Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std Dev
0.050	0.715	0.806	0.926	1.000	0.211	0.158

2b7.3. What is your interpretation of the results in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias**?** (i*.e., what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis, provide rationale for the selected approach for missing data*)

We do not believe any biases are introduced in the assessing of individual physician performance and continued endorsement of this measure would lead to improved care.

3. Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. **Data Elements Generated as Byproduct of Care Processes.**

Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. **To what extent are the specified data elements available electronically in defined fields?** *(i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields)*

ALL data elements are in defined fields in electronic clinical data (e.g., clinical registry, nursing home MDS, home health OASIS)

3b.2. **If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources.**

3b.3. **If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL.**

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.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 and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues.**

IF a PRO-PM, consider implications for both individuals providing PROM data (patients, service recipients, respondents) and those whose performance is being measured.

We have not identified any areas of concern or made any modifications as a result of testing and operational use of the measures in relation to data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, and other feasibility issues unless otherwise noted.

3c.2. **Describe any fees, licensing, or other requirements to use any aspect of the measure as specified** *(e.g., value/code set, risk model, programming code, algorithm)*.
Please seek licensing agreement prior to use of this measure.

4. Usability and Use

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. **Current and Planned Use**
NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement in addition to performance improvement.

Planned	Current Use (for current use provide URL)
	Quality Improvement (Internal to the specific organization) PINNACLE Registry http://www.ncdr.com/webncdr/pinnacle

4a.1. **For each CURRENT use, checked above, provide:**

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included

PINNACLE Registry (URL: http://www.ncdr.com/webncdr/pinnacle/)
In 2008, the American College of Cardiology Foundation launched the PINNACLE program (Formerly known as the Improving Continuous Cardiac Care or IC3). The PINNACLE Registry® continues to grow rapidly, with more than 5,700 providers representing over 1,500 unique office locations across the U.S submitting data to the registry. As of the fourth quarter of 2014, the registry has more than 28 million patient encounter records. PINNACLE assists practices in understanding and improving care through the production and distribution of quarterly performance reports. These reports, covering all valid patient encounters, detail adherence to 22 cardiovascular clinical measures at the physician, location, and practice levels across coronary artery disease, hypertension, heart failure and atrial fibrillation.

4a.2. **If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons?** (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)
We are continuously seeking opportunities to advocate for expanded use of this measure in government or other programs, including those intended for accountability or public reporting. The ACC and AHA do not have any policies that would restrict access to the performance measure specifications or results or that would impede implementation of the measure for any application. We would welcome its implementation in emerging applications such as accountable care organizations (ACO), Medicare Advantage insurance plans or health plans selling on the new insurance marketplace.

4a.3. **If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement.** (*Credible plan includes the specific program, purpose, intended audience, and timeline for*

implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.)
As described above, it is our understanding that CMS is also planning to move towards publicly reporting physician data via Physician Compare.

We are continuously seeking opportunities to advocate for expanded use of this measure in government or other programs, including those intended for accountability or public reporting such as PQRS reporting (including QCDR).

4b. Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b.1. **Progress on Improvement. (Not required for initial endorsement unless available.)**
Performance results on this measure (current and over time) should be provided in 1b.2 and 1b.4. Discuss:

- **Progress (trends in performance results, number and percentage of people receiving high-quality healthcare)**
- **Geographic area and number and percentage of accountable entities and patients included**

The mean performance rates from the Pinnacle Registry increased from 2013 to 2014 from 16.3 to 20.9%. The number of participating providers and patients also increased from 2013 to 2014. In 2013, 1701 providers were measured, and the patient study sample equals 209,770 . In 2014, 1,890 providers were measured, and the patient study sample equaled 239,948. The statistical significance of these results was not analyzed.

While the ACC and AHA create measures with an ultimate goal of improving the quality of care, measurement is a mechanism to drive improvement but does not equate with improvement. Measurement can help identify opportunities for improvement with actual improvement requiring making changes to health care processes and structure. In order to promote improvement, quality measurement systems need to provide feedback to front-line clinical staff in as close to real time as possible and at the point of care whenever possible. (1)

1. Conway PH, Mostashari F, Clancy C. The future of quality measurement for improvement and accountability. JAMA. 2013 Jun 5;309(21):2215-6.

4b.2. **If no improvement was demonstrated, what are the reasons? If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.**
Not Applicable.

4c. Unintended Consequences

The benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

4c.1. **Were any unintended negative consequences to individuals or populations identified during testing; OR has evidence of unintended negative consequences to individuals or populations been reported since implementation? If so, identify the negative unintended consequences and describe how benefits outweigh them or actions taken to mitigate them.**
We are not aware of any unintended consequences at this time, but we continuously monitor for them.

5. Comparison to Related or Competing Measures

If a measure meets the above criteria and there are endorsed or new related measures (either the same measure focus or the same target population) or competing measures (both the same measure focus and the same target population), the measures are compared to address harmonization and/or selection of the best measure.

5. Relation to Other NQF-endorsed Measures
Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

Yes

5.1a. **List of related or competing measures (selected from NQF-endorsed measures)**
0074 : Chronic Stable Coronary Artery Disease: Lipid Control
0118 : Anti-Lipid Treatment Discharge
0439 : STK-06: Discharged on Statin Medication
1519 : Statin Therapy at Discharge after Lower Extremity Bypass (LEB)

5.1b. **If related or competing measures are not NQF endorsed please indicate measure title and steward.**
The below are NQF endorsed measures but do not appear in the drop down for 5.1a.
• 696: STS CABG Composite Score (STS)
• 964: Therapy with aspirin, P2Y12 inhibitor, and statin at discharge following PCI in eligible patients (ACC)
• 2452: Percutaneous Coronary Intervention (PCI): Post-procedural Optimal Medical Therapy (ACC)

5a. Harmonization

The measure specifications are harmonized with related measures;
OR
The differences in specifications are justified

5a.1. **If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):**
Are the measure specifications completely harmonized?
No

5a.2. **If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.**
See discussion under 5b.1

5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. **If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):**
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)
This new measure on statin therapy for patients with ASCVD is based on the 2013 ACC/AHA guidelines, which focus on optimal treatment of statins. Most measures on statin therapy that are NQF-endorsed address subsets of the patients included in this broad measure and do not yet reflect the updated recommendations and/or are intended to be used in a different setting, level of analysis or different data source. Specific comments for each measure are below:

- 964: Therapy with aspirin, P2Y12 inhibitor, and statin at discharge following PCI in eligible patients (ACC)
- 2452: Percutaneous Coronary Intervention (PCI): Post-procedural Optimal Medical Therapy (ACC)

Measures 118 and 696 are STS registry-based measures and Measure 1519 is a SVS registry-based measure; thus, the data source and level of analysis are the same across all of the four measures. These measures are intended to be used at the time of hospital discharge, which differs from this new measure. In addition, based on the information provided on QPS, these measures do not reflect the updated recommendations for statin therapy.

Measure 439 is similar to the three measures discussed above with the exception of data source (electronic clinical data, paper medical records) and level of analysis (hospital/acute care facility). Similar concerns with the lack of alignment with the new ACC/AHA guidelines exist.

ACC/AHA believe that this new measure should be considered superior as it is aligned with the current recommendations and underlying evidence and is broadly applicable.

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

Attachment **Attachment:** ASCVDStatinAppendixA1.pdf

Contact Information

Co.1 **Measure Steward (Intellectual Property Owner):** American College of Cardiology
Co.2 **Point of Contact:** Penelope, Solis, comment@acc.org, 202-375-6576-
Co.3 **Measure Developer if different from Measure Steward:** American College of Cardiology
Co.4 **Point of Contact:** Penelope, Solis, comment@acc.org, 202-375-6576-

Additional Information

Ad.1 **Workgroup/Expert Panel involved in measure development**
Provide a list of sponsoring organizations and workgroup/panel members' names and organizations. Describe the members' role in measure development.
Members of the expert work group that developed the measure included: Joseph P. Drozda, Jr., MD, FACC, Chair (Mercy Health Director of Outcomes Research); T. Bruce Ferguson, Jr., MD, FACC, FAHA (Brody School of Medicine at ECU, Department of Cardiovascular Sciences); Brahmajee K. Nallamothu, MD, FACC (University of Michigan—Assistant Professor, Internal Medicine, Division of Cardiology); Hani Jneid, MD, FACC, FAHA, FSCAI (Baylor College of Medicine- MEDVAMC); Jeffrey W. Olin, DO, FACC, FAHA, MSVM (Mt. Sinai School of Medicine); Harlan M. Krumholz, MD, SM, FACC (Yale University School of Medicine); Henry H. Ting, MD, MBA, FACC, FAHA (New York-Presbyterian Hospital/Columbia University Medical Center).

ACCF, AHA, measures are developed through cross-specialty, multi-disciplinary work groups. All medical specialties and other health care professional disciplines participating in patient care for the clinical condition or topic under study are invited to participate as equal contributors to the measure development process.

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.2 **Year the measure was first released:** 2015
Ad.3 **Month and Year of most recent revision:** 12, 2015
Ad.4 **What is your frequency for review/update of this measure?** Specifications updated occur annual. See additional information section for more details.
Ad.5 **When is the next scheduled review/update for this measure?** 12, 2017

Ad.8 **Additional Information/Comments:** The ACCF, AHA have a formal measurement review process that stipulates regular (usually on a three-year cycle, when feasible) review of the measures. The process can also be activated if there is a major change in scientific evidence, results from testing or other implementation issues are noted that materially affect the integrity of the measure.