NATIONAL QUALITY FORUM

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 #: 0067

Corresponding Measures:

De.2. Measure Title: Coronary Artery Disease (CAD): Antiplatelet Therapy

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 (CAD) seen within a 12-month period who were prescribed aspirin or clopidogrel

1b.1. Developer Rationale: In the absence of contraindications, antiplatelet therapy is recommended for patients with a diagnosis of coronary artery disease as it reduces the risk of adverse coronary events, including death, by inhibiting platelet aggregation. Aspirin and clopidogrel have been proven to stabilize coronary plaque and prevent ruptures and clots. Despite the strong evidence, adherence to these guideline-directed therapies is suboptimal.

S.4. Numerator Statement: Patients who were prescribed aspirin or clopidogrel

S.6. Denominator Statement: All patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12-month period

S.8. Denominator Exclusions: Denominator exceptions

- Documentation of medical reason(s) for not prescribing aspirin or clopidogrel (e.g., allergy, intolerance, receiving other thienopyridine therapy, receiving warfarin therapy, bleeding coagulation disorders, other medical reasons)
- Documentation of patient reason(s) for not prescribing aspirin or clopidogrel (e.g., patient declined, other patient reasons)
- Documentation of system reason(s) for not prescribing aspirin or clopidogrel (e.g., lack of drug availability, other reasons attributable to the health care system)

De.1. Measure Type: Process

S.17. Data Source: Registry Data

S.20. Level of Analysis: Clinician : Individual

IF Endorsement Maintenance – Original Endorsement Date: Aug 10, 2009 **Most Recent Endorsement Date:** Feb 19, 2016

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results?

Preliminary Analysis: Maintenance of Endorsement

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 *structure, 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. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

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 2016

- The developer provided decision logic from secondary prevention to outcome for the use of antiplatelet therapy in decreasing morbidity, mortality, and hospitalization with patients in with chronic stable CAD.
- The developer provided two guidelines with four guideline statements supporting the use of aspirin and clopidogrel in patients with CAD:
    - 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable IHD
        - Non–enteric-coated, chewable aspirin (162 mg to 325 mg) should be given to all patients with NSTE-ACS without contraindications as soon as possible after presentation, and a maintenance dose of aspirin (81 mg/d to 325 mg/d) should be continued indefinitely. **Class I: Level of Evidence: A**
        - In patients with NSTE-ACS who are unable to take aspirin because of hypersensitivity or major gastrointestinal intolerance, a loading dose of clopidogrel followed by a daily maintenance dose should be administered. **Class I: Level of Evidence: B**

    - 2014 AHA/ACC guideline for the management of patients with non–ST-elevation acute coronary syndromes
        - Treatment with aspirin 75 to 162 mg daily should be continued indefinitely in the absence of contraindications in patients with SIHD. **Class I: Level of Evidence: A**
        - Treatment with clopidogrel is reasonable when aspirin is contraindicated in patients with SIHD. **Class I: Level of Evidence: B**
- In 2016, the developer had cited one meta-analysis for STEMI patients, which was published after the publication of the 2012 discussed guideline comparing intravenous P2Y12 inhibitors with clopidogrel on major ischemic and bleeding events, though the developer stated that it does not conflict with the 10 guideline recommendation statements.
    - In this 2020 submission, the developer included updated literature search covering January 1, 2016 through February 19, 2020 for STEMI patients, and confirmed that none of the studies contained new conclusions that would alter the recommendation to prescribe antiplatelet therapy to patients with coronary artery disease.
    - The developer also included a summary of the Quantity, Quality and Consistency of the body of 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:**

Questions for the Committee:

- *The developer attests the underlying evidence for the measure has not changed since the last NQF endorsement review in 2016. Does the Committee agree the evidence basis for the measure has not changed and there is no need for repeat discussion on Evidence?*
- *Does the Committee agree to accept the rating from previous year's evaluation and not re-vote on Evidence?*

Guidance from the Evidence Algorithm

Process measure based on systematic review (Box 3) → QQC presented (Box 4) → Quantity: high; Quality: high; Consistency: high (Box 5) → High (Box 5a)

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 January 2018-December 2018 data from 1,846 providers who reported on this measure through the registry reporting for MIPS. The dataset reflects information at the provider level.
    - Of those 1,846 providers, 1,843 providers had at least one patient who qualified for the measure, after accounting for exceptions, for a total of 506,259 eligible patients.

    - The average number of eligible patients is 274 for the 1,843 providers.
    - The range of eligible patients for 1,843 providers is from 1 to 2,781
- Based on the sample of 1,843 included providers, the developers reported:
    - The mean performance rate of 0.88
    - The median performance rate of 0.88
    - The mode of 1.0
    - The standard deviation of 0.19
    - The range of the performance rate of 1.00, with a minimum rate of 0.00 and a maximum rate of 1.00.
    - The interquartile range of 0.15 (1.00–0.85)
- The developers also reported CMS published quality benchmarks for MIPS 2020, 2019 and 2018, which are created using historical performance rates, which included 2,407 providers, and the patient study sample of 1,023,530. The data demonstrated:
    - Overall mean performance on this measure was 86.2%, with a standard deviation of 10.5%. The minimum score equaled 0.00%, while the maximum score equaled 100.00%. The interquartile score equaled to 10.3%

Year	Submission Method	Average Performance Rate	Standard Deviation
2018	CQM	89.2	N/A
2017	Registry/QCDR	89.6	13.2
2016	Registry/QCDR	87.3	11.72013 performance data from the Pinnacle registry.

- The developers also reported 2014 performance data from the Pinnacle registry, which included 2,248 providers, and the patient study sample of 959,792. The data demonstrated:
    - Overall mean performance of 86.3%, with a standard deviation of 9.49%. The minimum score of 0.00%, while the maximum score of 100.00%. The interquartile score equaled to 10.2%.

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 Tran et al. (2017) using data from National Health and Nutrition Examination Survey. The data found that among those with coronary artery disease (CAD), women and racial/ethnic minorities were less likely to take aspirin, compared with men and non-Hispanic Whites (OR = 0.63 and 0.74 respectively)

Questions for the Committee:

- Does the performance data provided continue to warrant a national performance measure?
- 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

- Strong evidence level (1A)
- Evidence supports measure
- no comments
- Existing measure. There is a systematic review of the evidence: 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the Diagnosis and Management of Patients with Stable Ischemic Heart Disease that provides information on the quantity, quality and consistency. Included updated literature search that yielded no data to change current recommendations or measure. Agree with preliminary rating as High
- Strong evidence. High rating.
- Similar to the last measure, there is a focused update to the 2014 AHA/ACC guideline.
- Strong evidence, but would reconsider re-titling for secondary prevention.
- high
- Evidence is high. No need to discuss.
- no concerns about evidence

1b. Performance Gap

- Mean/median scores in upper 80%s suggesting some but not a lot of room to improve. No data on racial/ethnic disparities given.
- No performance gaps
- no comments
- 2018 MIPS data show range 0 - 1.0 with mean of 0.88 and mode of 1.0 There is only modest room for improvement overall. There is no data from performance monitoring to identify disparities. Developers cite the Tran et al. article from 2017 using NHANES data that shows women and racial/ethnic minorities were less likely to take ASA than men or non-Hispanic Whites
- Some performance gap but not that large.
- There is a high level of performance on this measure. It can still help improve quality of care delivered, but would recommend looking into disparities and whether the measure should be risk adjusted.
- Appears to be more a socio-economic performance gap than provider gap.
- moderate
- Moderate: while performance is good, there are racial and gender disparities.
- it's beginning to top out, limited room for improvement

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

2b. Validity: Testing; Exclusions; Risk-Adjustment; Meaningful Differences; Comparability; Missing Data

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

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 – less emphasis if no new testing data provided.

Validity

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 – less emphasis if no new testing data provided.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Composite measures only:

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.

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: NQF Staff

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- The staff is satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss and/or vote on reliability?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- The staff is satisfied with the validity analyses for the measure. Does the Committee think there is a need to discuss and/or vote on validity?

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

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

Staff Evaluation of Scientific Acceptability

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 0067

Measure Title: Coronary Artery Disease (CAD): Antiplatelet Therapy

Type of measure:

☒ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☐ **Claims** ☐ **Electronic Health Data** ☐ **Electronic Health Records** ☐ **Management Data**
☐ **Assessment Data** ☐ **Paper Medical Records** ☐ **Instrument-Based Data** ☒ **Registry Data**
☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☐ **Clinician: Group/Practice** ☒ **Clinician: Individual** ☐ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐ **New** ☒ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

   Submission document: "MIF_0067" document, items S.1-S.22

   ***NOTE**: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**
   - No concerns

RELIABILITY: TESTING

Submission document: "MIF_0067" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☐ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical <u>VALIDITY</u> testing** of <u>patient-level data</u> conducted?

   ☐ **Yes** ☐ **No**

6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

- Reliability testing was conducted on the performance score, using registry data from the 2018 Merit-based Incentive Payment System (MIPS) Program for the time period of January 1st, 2018 thought December 31st, 2018; and the data analysis only included data that was reported at the unique NPI level.
    - A total of 1,846 providers are reporting on this measure through the registry reporting option for MIPS. Of those 1,846 providers, 1,843 providers had at least one patient who qualified for the measure, after accounting for exceptions, for a total of 506,259 eligible patients. The average number of eligible patients is 274 for the 1,843 providers. The range of eligible patients for 1,843 providers is from 1 to 2,781.
- The developers used a beta-binomial model to assess the signal-to-noise ratio to conduct reliability testing. 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.

7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

   - The average reliability for providers with at least one eligible patient is 0.95. As the developer states, 0.80 – 0.90 is considered high reliability.
   - The developer also reported the average reliability at each decile of the sample

8. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? NOTE: If multiple methods used, at least one must be appropriate.

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

   ☐ **Not applicable** (score-level testing was not performed)

9. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

   ☒ **Not applicable** (data element testing was not performed)

10. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

    ☒ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

    ☐ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has not been conducted)

    ☐ **Low** (NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

    ☐ **Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)

11. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

    Precise, unambiguous and complete specification (Box 1) → reliability testing conducted with computed measure scores for each measured entity (Box 4) → based on reliability statistic and scope of testing, there is a high confidence that the performance measure scores are reliable (Box 6a) → High

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

12. **Please describe any concerns you have with measure exclusions.**

    Submission document: Testing attachment, section 2b2.

    - The developer did not include exclusions and does specify the exceptions that were analyzed.

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

    Submission document: Testing attachment, section 2b4.

    - No concerns

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
    Submission document: Testing attachment, section 2b5.

    - N/A

15. **Please describe any concerns you have regarding missing data.**

    Submission document: Testing attachment, section 2b6.

    - The developer stated that the MIPS dataset provided to us by CMS did not contain missing data, so this test was not performed. Nevertheless, missing data may have been rejected when submitted to CMS in which case those values would not be counted towards measure performance. There is no indication that this missing data was systematic, thus their omission would lead to unbiased performance results.

16. **Risk Adjustment**

    16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

    16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

    ☐ Yes ☐ No ☒ Not applicable

    16c. **Social risk adjustment:**

    16c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☒ Not applicable

    16c.2 Conceptual rationale for social risk factors included? ☐ Yes ☐ No

    16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☐ No

    16d.**Risk adjustment summary:**

    16d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

    16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

    16d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

    16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

    16d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

16e**. Assess the risk-adjustment approach**

VALIDITY: TESTING

17. **Validity testing level: ☒ Measure score ☐ Data element ☐ Both**
18. **Method of establishing validity of the measure score:**
    - ☐ **Face validity**
    - ☒ **Empirical validity testing of the measure score**
    - ☐ **N/A (score-level testing not conducted)**
19. **Assess the method(s) for establishing validity**

    Submission document: Testing attachment, section 2b2.2

    - The developers used data from the 2018 MIPS Registry Program to perform the correlation analysis for this measure. Data came from the Registry version of Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) (NQF 0066) and Coronary Artery Disease (CAD): Antiplatelet Therapy (NQF 0067)
    - The developers chose Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) (NQF 0066) to conduct correlation analysis due to the similarities in patient population and domain. They hypothesized that there exists a positive association of scores between providers who prescribe an antiplatelet therapy on patients with coronary artery disease and those who prescribe Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy on patients with diabetes or left ventricular systolic dysfunction (LVEF < 40%) within a 12 month period.
    - Providers included in the analysis had at least one patient in the denominator after exceptions were removed. Results identify the multiple R value (the correlation coefficient) and P-value of the regression variables to assess the association between performance scores of these shared provider IDs.
20. **Assess the results(s) for establishing validity**

    Submission document: Testing attachment, section 2b2.3

    - The results demonstrated that the Coronary Artery Disease (CAD): Antiplatelet Therapy (NQF 0067) was positively correlated with Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) (NQF 0066).
    - The correlation was highly statistically significant with a coefficient of correlation of 0.47, which showed moderate correlation, significant, and confirms the developer hypothesis. The moderate positive correlation with Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) (NQF 0066) demonstrates the criterion validity of the measure.
21. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

    Submission document: Testing attachment, section 2b1.

    ☒ **Yes**

    ☐ **No**

☐ **Not applicable** (score-level testing was not performed)

22. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?**
NOTE that data element validation from the literature is acceptable.

Submission document: *Testing attachment, section 2b1.*

☒ **Yes**

☐ **No**

☐ **Not applicable** (data element testing was not performed)

23. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

☐ **Low** (NOTE: Should rate LOW if you believe that there are threats to validity and/or relevant threats to validity were not assessed OR if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level is required; if not conducted, should rate as INSUFFICIENT.)

24. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

All potential threats to validity are empirically assessed (Box 1) → Validity testing conducted with computed measure scores (Box 5) → Validity testing method was described and appropriate for assessing hypothesized relationships (Box 6) → moderate confidence that the performance measure scores are a valid indicator of quality (Box 6b) → Moderate

ADDITIONAL RECOMMENDATIONS

25. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a. Reliability

- No issues
- no concerns
- no comments
- No issues regarding data elements, logic etc. Existing measure first endorsed 2009. Most recent endorsement 2016
- Clear specifications of previously endorsed measure.
- measure had high reliability even with a low sample size.
- Inclusion criteria do not specify beyond CAD. Subclinical CAD may be diagnosed/coded more than previously due to more use of imaging and risk-adjusted coding. Sub-clinical CAD is not necessarily appropriate for anti-platelet therapy.

- high
- Reliability is high. The measure has been used for 11 years. No need to discuss.
- no concerns No issues
- no
- none
- Signal to noise testing: Mean 0.95 - very good but interestingly less than 2013 & 2014
- Very good reliability. Beta-binomial model used with average reliability of 0.95. High.
- aligned with nqf assessment
- See above comment
- no
- No concerns
- No

2b. Validity -Testing

- Moderate IMO. Not sure that correlation with performance on a CHF measure constitutes a robust validity test.
- no
- none
- Empiric validity - correlation with ACEI/ARB in CAD with DM or decreased LVEF measure performance 0.47 No concerns
- Moderate validity with correlation of 0.47.
- I appreciate even though moderate positive correlation is it statistcally significant. i put more emphasis on significance, because measures may never be perfectly aligned.
- No
- no
- No particular concerns. tested agains ACEi prescription.
- No
- N/A
- no
- no comments
- MIPS data provided by CMS had no missing data. No evidence of systematic omissions
- None.
- no concerns and well documented by developer
- Aspirin use may not be always captured as a prescription drug, however this should not affect relative performance on the measure.
- no
- No significant threats to validity. Some of the low performance might be due to not getting ASA on the med list. (Sloppy)
- no concerns
- No risk adjustment done, which can compromise validity of results.

- yes
- no comment
- There is no risk adjustment. Exclusions are appropriate.
- No risk adjustment.
- would like to see some risk adjustment be considered for future evaluation
- Exclusions are appropriate
- no
- This is not risk adjusted. No other threats
- no concerns

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 developer noted that the data elements are generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score) and are coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)
- ALL data elements are in defined fields in electronic clinical data (e.g., clinical registry, nursing home MDS, home health OASIS).
- All the data elements needed for this measure are collected through electronic data

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?
- If an eCQM, does the eCQM 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

- No concerns. Already part of MIPS
- none
- no comments
- Demonstrated feasibility - existing measure with years of performance data
- Moderate.

- No major concerns. Seems appropriate set of data for level of analysis
- No issues.
- high
- All of the data elements are collected during routine care.Feasibility is moderate
- no concerns

Criterion 4: Usability and Use

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

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. 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.

4a.1. 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.

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

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

Accountability program details

- Performance results are used in the following programs:
    - Merit-based Incentive Payment System MIPS) – sponsored by the Centers for Medicare and Medicaid Services (CMS) with the purpose to tie payments to clinicians, to provide quality and cost-efficient care, drive improvement in care processes, increase the use of healthcare data, and reduce the costs of care. The program began in 2017, nationally, and the clinicians are included in MIPS-Quality if they are an eligible clinician type and meet program requirements.
    - The PINNACLE Registry – sponsored by American College of Cardiology and its National Cardiovascular Data Registry. It began in 2008 and it was developed as an outpatient-based prospective quality improvement registry, for cardiology practices in the outpatient setting. Data is collected specifically for coronary artery disease, hypertension, heart failure and atrial fibrillation. Data collected includes patient demographics, medical history, vital signs, laboratory values, imaging results, medications, and contraindications to medications. It's implemented nationally.

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given

an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- Performance results are provided to all participating clinicians across quality programs. Clinicians participating in the MIPS-Quality program receive performance results, i.e. their MIPS score and corresponding payment adjustment annually. The PINNACLE Registry provides user-friendly online benchmark reports to users via an interactive portal. Individual practice performance on measures is provided by quarter, alongside the national average by quarter. This allows clinicians to identify areas for improvement. The Diabetes Collaborative Registry provides user-friendly online benchmark reports.
- The feedback process includes providing user-friendly online benchmark reports to users via an interactive portal. Individual practice performance on measures is provided by quarter, alongside the national average by quarter.

Additional Feedback:

- None

***Questions for the Committee**:*

- How have (or can) the performance results be used to further the goal of high-quality, efficient healthcare?
- How has the measure been vetted in real-world settings by those being measured or others?

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

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability 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.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- The developers noted that the MIPS quality benchmarks of this form show a slight improvement in average performance rates between 2016 and 2018. The scores appear very similar.
- Developers also cited data from the literature: Makam et al. (2016) evaluated trends in the prescription of evidence-based cardiac medications amongst eligible patients. Methods included analyzing over 5,000 patients discharged from various hospitals in central Massachusetts after acute myocardial infarction. The study found that eligible patients receiving a new prescription for aspirin increased from 70% to 88% between 2001 and 2011.

4b2. Benefits vs. harms. 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).

Unexpected findings (positive or negative) during implementation

- The developer provided data from a 2017 meta-analysis by Du, Cheng, Zhang, Li, & Mei that evaluated the relationship between medication adherence and clinical outcomes amongst ~100,000 patients with stable coronary artery disease.
- Study authors found that adherence to evidence-based medications, including antiplatelet drugs, was related to a lower risk of all-cause mortality (risk ratio 0.56; 95% CI), lower risk of cardiovascular mortality (risk ratio 0.66; 95% CI), and lower risk of cardiovascular hospitalization/myocardial infarction(risk ratio 0.61; 95% CI)

Potential harms

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

Additional Feedback:

- N/A

***Questions for the Committee**:*

- Based on the improvement data provided, does the Standing Committee agree that there has been improvement in average performance rate?
- 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. Use

- no issues
- yes
- no comments
- Publicly reported & used in accountability programs - MIPS, PINNACLE; Developer reports feedback through Quality Measures help desk, public comment and ACC/AHA Task Force on Performance Measures. No feedback leading to measure change
- Pass. MIPS and one registry.
- In two major measurement systems (MIPS, PINNACLE).
- Inter-provider variation is narrowing, but there is parallell evidence that some patient groups are still lagging in performance.
- high
- The measure is publicly reported and used in accountability programs
- no concerns

4b. Usability

- None
- none
- no comments

- Based on the MIPS data provided there does not appear to be any meaningful improvement. Performance is linked to guideline adherence which would be an indicator of quality care. Benefits outweigh the risks although we are learning more about GIB risk all the time especially in the elderly.
- Moderate to high.
- aligned with NQF
- Usable, no harms
- high
- Benefit outweighs harm in this population
- no concerns

Criterion 5: Related and Competing Measures

Related or competing measures

- 0465 : Perioperative Anti-platelet Therapy for Patients undergoing Carotid Endarterectomy

NQF Staff identified additional related measures:
0073: Ischemic Vascular Disease (IVD): Use of Aspirin of Another Antiplatelet
0076: Optimal Vascular Care

Harmonization

- The developer stated that the specifications are harmonized to the extent possible; The patient population of 0465 is adults undergoing carotid endarterectomy, whereas the patient population of 067 is adults with coronary artery disease.

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

No

- yes
- no
- related measure 0465 Antiplatelet Therapy for Carotid Endarterectomy - while both are atherosclerotic diseases the population and intent is different
- Related measures but harmonized.
- would like to hear how this measure differs from the NQF identified measures.
- Is it harmonized with 0073? There will be a lot of overlap with patients in both measures.
- no
- 0465 is related buy not competing
- harmonized to the degree possible

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: June 12, 2020

- **No NQF Members have submitted support/non-support choices as of this date.**
- **No Public or NQF Member comments submitted as of this date.**

Developer Submission

1. Evidence and Performance Gap – 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 sub criteria to pass this criterion and be evaluated against the remaining criteria.***

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

2020_NQF_evidence_attachment_0067_Final.docx

1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

No

Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0067

Measure Title: Coronary Artery Disease: Antiplatelet Therapy

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: 4/9/2020

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

Outcome

☐ 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. (*A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)*

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

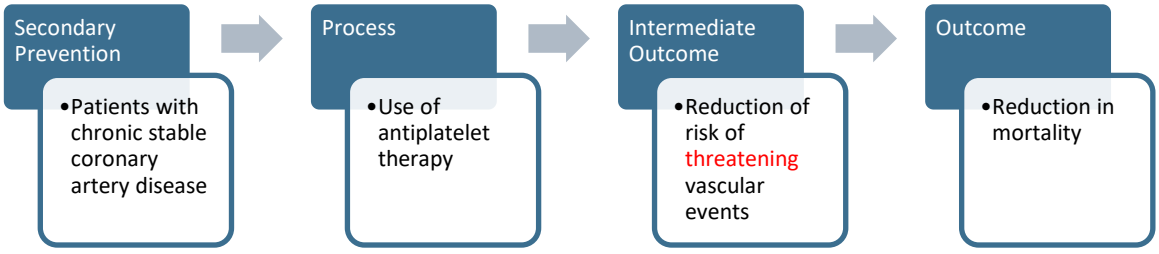
☒ Process: Antiplatelet therapy for patients with coronary artery disease

☐ Appropriate use measure: Click here to name what is being measured

☐ Structure: Click here to name the structure

☐ Composite: Click here to name what is being measured

1a.2 LOGIC MODEL Diagram or briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.



Use of antiplatelet therapy can decrease morbidity, mortality, and hospitalizations for patients with chronic stable coronary artery disease.

1a.3 Value and Meaningfulness: **IF** this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

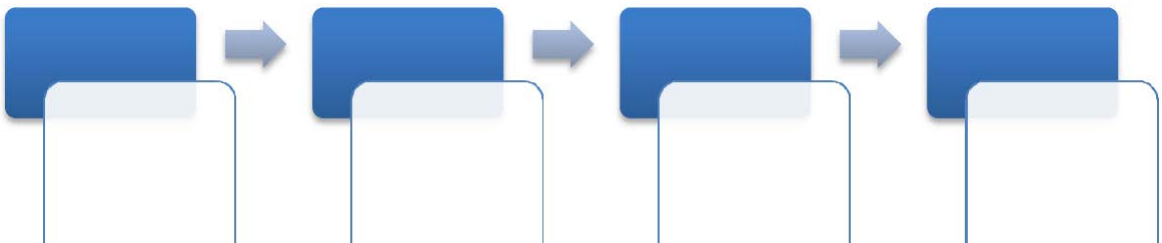
What is the source of the systematic review of the body of evidence that supports the performance measure? A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data. (IOM)

X Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other



Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	**The guidelines supporting this measure remain unchanged since the last submission. However, we included additional detail from a second set of guidelines. The American Heart Association (AHA) notes that evidence is reviewed at least twice a year, and updates are initiated on an as-needed basis.** **2014 Guidelines** **Title: 2014 AHA/ACC guideline for the management of patients with non–ST-elevation acute coronary syndromes: a report of the American College of Cardiology/ American Heart Association Task Force on Practice Guidelines. A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines** **Author: Ezra A. Amsterdam, Nanette K. Wenger, Ralph G. Brindis, Donald E. CaseyJr, Theodore G. Ganiats, David R. HolmesJr, Allan S. Jaffe, Hani Jneid, Rosemary F. Kelly, Michael C. Kontos, Glenn N. Levine, Philip R. Liebson, Debabrata Mukherjee, Eric D. Peterson, Marc S. Sabatine, Richard W. Smalling, and Susan J. Zieman** **Date: 2014** **Citation, including page number: Amsterdam, E. A., Wenger, N. K., Brindis, R. G., Casey, D. E., Jr, Ganiats, T. G., Holmes, D. R., Jr, … Zieman, S. J. (2014). 2014 AHA/ACC Guideline for the Management of Patients with Non-ST-Elevation Acute Coronary Syndromes: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Journal of the American College of Cardiology, 64*(24), e161. doi:10.1016/j.jacc.2014.09.017** **URL: https://www.ahajournals.org/doi/10.1161/CIR.0000000000000134** **2012 Guidelines** • **Title: 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.** • **Author: Stephan D. Fihn, Julius M. Gardin, Jonathan Abrams, Kathleen Berra, James C. Blankenship, Apostolos P. Dallas,**

	Pamela S. Douglas, JoAnne M. Foody, Thomas C. Gerber, Alan L. Hinderliter, Spencer B. KingIII, Paul D. Kligfield, Harlan M. Krumholz, Raymond Y.K. Kwong, Michael J. Lim, Jane A. Linderbaum, Michael J. Mack, Mark A. Munger, Richard L. Prager, Joseph F. Sabik, Leslee J. Shaw, Joanna D. Sikkema, Craig R. SmithJr, Sidney C. SmithJr, John A. Spertus, and Sankey V. Williams • **Date: 2012** • **Citation: Fihn, S. D., Gardin, J. M., Abrams, J., Berra, K., Blankenship, J. C., Dallas, A. P., … Society of Thoracic Surgeons (2012). 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. *Journal of the American College of Cardiology, 60*(24), page 50. doi:10.1016/j.jacc.2012.07.013** • **URL: https://www.ahajournals.org/lookup/doi/10.1161/CIR.0b013e318277d6a0**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	2014 Guidelines 1. Non–enteric-coated, chewable aspirin (162 mg to 325 mg) should be given to all patients with NSTE-ACS without contraindications as soon as possible after presentation, and a maintenance dose of aspirin (81 mg/d to 325 mg/d) should be continued indefinitely. **Class I: Level of Evidence: A** 2. In patients with NSTE-ACS who are unable to take aspirin because of hypersensitivity or major gastrointestinal intolerance, a loading dose of clopidogrel followed by a daily maintenance dose should be administered. **Class I: Level of Evidence: B** 2012 Guidelines 3. Treatment with aspirin 75 to 162 mg daily should be continued indefinitely in the absence of contraindications in patients with SIHD. **Class I: Level of Evidence: A** 4. Treatment with clopidogrel is reasonable when aspirin is contraindicated in patients with SIHD. **Class I: Level of Evidence: B**

<table>
<tr>
<td>Grade assigned to the evidence associated with the recommendation with the definition of the grade</td>
<td>
The Level of Evidence is an estimate of the certainty of precision of the treatment effect.

<u>2014 Guidelines:</u>
1. Non–enteric-coated, chewable aspirin (162 mg to 325 mg) I should be given to all patients with NSTE-ACS without contraindications as soon as possible after presentation, and a maintenance dose of aspirin (81 mg/d to 325 mg/d) should be continued indefinitely. Level of Evidence: A
2. In patients with NSTE-ACS who are unable to take aspirin because of hypersensitivity or major gastrointestinal intolerance, a loading dose of clopidogrel followed by a daily maintenance dose should be administered. Level of Evidence: B

<u>2012 Guidelines</u>
3. Treatment with aspirin 75 to 162 mg daily should be continued indefinitely in the absence of contraindications in patients with SIHD. Level of Evidence: A
4. Treatment with clopidogrel is reasonable when aspirin is contraindicated in patients with SIHD. Level of Evidence: B

<table>
<tr><td rowspan="2">Level A:</td><td>Multiple populations evaluated*</td></tr>
<tr><td>Data derived from multiple randomized clinical trials or meta-analyses</td></tr>
<tr><td rowspan="2">Level B:</td><td>Limited populations evaluated*</td></tr>
<tr><td>Data derived from a single randomized trial or non-randomized studies</td></tr>
</table>
</td>
</tr>
<tr>
<td>Provide all other grades and definitions from the evidence grading system</td>
<td>
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.
</td>
</tr>
</table>

<table>
<tr>
<td></td>
<td>
Level of Evidence C: Consensus opinion of experts, case studies, or standard of care.

<table>
<tr><td rowspan="2">Level A:</td><td>Multiple populations evaluated*</td></tr>
<tr><td>Data derived from multiple randomized clinical trials or meta-analyses</td></tr>
<tr><td rowspan="2">Level B:</td><td>Limited populations evaluated*</td></tr>
<tr><td>Data derived from a single randomized trial or non-randomized studies</td></tr>
<tr><td rowspan="2">Level C:</td><td>Very limited populations evaluated*</td></tr>
<tr><td>Only consensus opinion of experts, case studies, or standard of care</td></tr>
</table>

*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.

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.
</td>
</tr>
<tr>
<td>Grade assigned to the recommendation with definition of the grade</td>
<td>
The Class of Recommendation is an estimate of the size of the treatment effect, with consideration given to risks versus benefits as well as evidence and/or agreement that a given treatment or procedure is or is not useful/effective or in some situations may cause harm.

<u>2014 Guidelines</u>
1. Non–enteric-coated, chewable aspirin (162 mg to 325 mg) I should be given to all patients with NSTE-ACS without contraindications as soon as possible after presentation, and a maintenance dose of aspirin (81 mg/d to 325 mg/d) should be continued indefinitely. Class I
2. In patients with NSTE-ACS who are unable to take aspirin because of hypersensitivity or major gastrointestinal
</td>
</tr>
</table>

<table>
<tr>
<td></td>
<td>
intolerance, a loading dose of clopidogrel followed by a daily maintenance dose should be administered. Class I

<u>2012 Guidelines</u>
3. Treatment with aspirin 75 to 162 mg daily should be continued indefinitely in the absence of contraindications in patients with SIHD. Class I
4. Treatment with clopidogrel is reasonable when aspirin is contraindicated in patients with SIHD. Class I

<table>
<tr><td rowspan="2">Class I</td><td>Benefit >>> Risk</td></tr>
<tr><td>Procedure/Treatment SHOULD be performed/administered</td></tr>
</table>
</td>
</tr>
<tr>
<td>Provide all other grades and definitions from the recommendation grading system</td>
<td>
<table>
<tr><td rowspan="2">Class I</td><td>Benefit >>> Risk</td></tr>
<tr><td>Procedure/Treatment SHOULD be performed/administered</td></tr>
<tr><td rowspan="2">Class IIa</td><td>Benefit >> Risk; Additional studies with focused objectives needed</td></tr>
<tr><td>IT IS REASONABLE to perform procedure / administer treatment</td></tr>
<tr><td rowspan="2">Class IIb</td><td>Benefit >= Risk; Additional studies with broad objectives needed; additional registry data would be helpful</td></tr>
<tr><td>Procedure / Treatment MAY BE CONSIDERED</td></tr>
<tr><td>Class III No Benefit</td><td>Procedure / Test = Not helpful
Treatment = No proven benefit</td></tr>
<tr><td>Class III Harm</td><td>Procedure / Test = Excess cost without benefit or harmful
Treatment = Harmful to patients</td></tr>
</table>

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

Recommendations with a Class I, Level B designation are characterized as:
</td>
</tr>
</table>

	-Recommendation that procedure or treatment is useful/effective -Evidence from single randomized trial or nonrandomized studies
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	2014 Guidelines: The cited body of evidence in support of the two recommendations includes 27 clinical trials. We analyzed the supporting studies and additional details are below: • Related meta-analyses that included 6 primary prevention trials, involving 95,000 individuals, and 16 secondary prevention trials, involving 17,000 patients comparing aspirin versus control. • Randomized controlled trial that assigned 12,562 patients to clopidogrel or a placebo. • Clinical trial involving 8560 patients assigned to double-dose and 8703 to standard-dose clopidogrel, and 8624 to high-dose and 8639 to low-dose aspirin. 30-day follow-up. • CAPRIE was a randomized, blinded, international trial and looked at clopidogrel and aspirin in reducing the risk of ischemic stroke, myocardial infarction, or vascular death. 2012 Guidelines: The body of evidence supporting the recommendations on antiplatelet therapy with patients with a prior MI includes randomized controlled trials and meta-analyses. The number of which is not provided in the guideline. All of the recommendations for this process are rated as Level of Evidence A or B, meaning that the data was derived from one or more RCTs or meta-analyses. Additional information on the overall quality of evidence across the RCTs is not provided. The cited evidence in support of the two recommendations include 1) a meta-analyses evaluating 287 studies that looked at outcomes of 135,000 patients receiving antiplatelet therapy versus control and outcomes of 77,000 patients receiving different antiplatelet regimens; 2) A randomized, double-blind clinical trial involving 2,035 patients receiving aspirin or placebo; 3) The CAPRIE trial described above. Both drugs are associated with significant benefits for patients with coronary artery disease.
Estimates of benefit and consistency across studies	2014 Guidelines

	The guidelines did not include an overall estimate of benefit and consistency from the body of evidence supporting the recommendations. We analyzed the supporting articles and found the following:
	• Related meta-analyses that included 6 primary prevention trials, involving 95,000 individuals, and 16 secondary prevention trials, involving 17,000 patients comparing aspirin versus control.
	  ○ Aspirin had a substantial net benefit for patients with occlusive vascular disease. In the secondary prevention trials, patients treated with aspirin had an absolute reduction in serious vascular events ($p<0.0001$).
	• Randomized controlled trial that assigned 12,562 patients to clopidogrel or a placebo.
	  ○ The clopidogrel group was less likely to suffer from death from cardiovascular causes, nonfatal myocardial infarction, or stroke (relative risk = 0.80) (CI = 91%) ($P<0.001$). The clopidogrel group was also less likely to suffer from refractory ischemia (relative risk = 0.86) ($P<0.001$).
	• Clinical trial involving 8560 patients assigned to double-dose and 8703 to standard-dose clopidogrel, and 8624 to high-dose and 8639 to low-dose aspirin. 30-day follow-up.
	  ○ Patients who received a double dose of clopidogrel had a lower risk of cardiovascular death, myocardial infarction, or stroke (adjusted hazard ratio = .86) ($P=0.039$). Patients on high-dose and low-dose aspirin had no difference in outcomes.
	
	In addition, according to the guidelines, the Class of Recommendation means the strength of the recommendation, which encompasses the anticipated magnitude and judged certainty of benefit in proportion to risk (Amsterdam et al., 2014). All four recommendations received a grade of Class I recommendation, which means that the benefits of the treatment, i.e. aspirin or clopidogrel for eligible patients, >>> the risks, and the treatment should be administered.
	
	<u>2012 Guidelines:</u>
	p. e95
	
	Among 2,920 patients with SIHD, a comprehensive meta-analysis of source data revealed an association of aspirin use with a 37% reduction in the risk of serious vascular events, including a 46% decrease in the risk for UA and a 53% decrease in the risk of requiring coronary angioplasty. (Collaborative meta-analysis of randomised trials of antiplatelet therapy for prevention of death,

	myocardial infarction, and stroke in high risk patients. BMJ. 2002;324:71– 86.) Clopidogrel 75 mg has been compared with aspirin 325 mg in patients with previous MI, stroke, or symptomatic PAD in the prospective, randomized CAPRIE (Clopidogrel versus Aspirin in Patients at Risk of Ischaemic Events) study. (A randomised, blinded, trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). CAPRIE Steering Committee. Lancet. 1996;348:1329 –39.) Although clopidogrel demonstrated superiority over aspirin in the secondary prevention of MI and death in this group of patients, the magnitude of difference was small. Because no additional trials comparing aspirin and clopidogrel in patients with SIHD have been conducted, clopidogrel remains an acceptable alternative agent to aspirin.
What harms were identified?	2014 Guidelines Study authors found that aspirin and clopidogrel increased the risk for bleeding. The CAPRIE study found that patients reported adverse side effects from clopidogrel and aspirin to include rash, diarrhea, upper gastrointestinal discomfort, intracranial hemorrhage, and gastrointestinal hemorrhage (1996). Another supporting study found that patients receiving clopidogrel were more likely to experience major bleeding, compared to the placebo group (relative risk = 1.38) (P=0.001). However, there were no cases of excess bleeding that caused strokes, required surgical intervention or inotropic agents, or caused permanent disability (Yusuf, et al., 2001). All four recommendations received a grade of Class I recommendation, which means that the benefits of the treatment, i.e. aspirin or clopidogrel for eligible patients, >>> the risk of administering this treatment, and the treatment should be administered. 2012 Guidelines 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease The guideline does not discuss potential harms of aspirin or clopidogrel therapy alone but the potential risks for major bleeding when both therapies are prescribed is discussed. p. e95

	In a meta-analysis of 5 RCTs comparing clopidogrel plus aspirin to aspirin alone in patients with IHD, the incidence of all-cause mortality, MI, and stroke was found to be reduced in the clopidogrel-plus- aspirin group, whereas the risk of major bleeding increased significantly. The incidence of all-cause mortality was 6.3% in the aspirin plus clopidogrel group versus 6.7% in the aspirin group (odds ratio [OR] 0.94; 95% CI 0.89, 0.99; p = 0.026). The incidence of myocardial infarction was 2.7% and 3.3% (OR 0.82; 95% CI 0.75, 0.89; p < 0.0001), and stroke was 1.2% and 1.4% (OR 0.82; 95% CI 0.73, 0.93; p = 0.002). Similarly, the incidence of major bleeding was 1.6% and 1.3% (OR 1.26; 95% CI 1.11, 1.41; p < 0.0001), and fatal bleeding was 0.28% and 0.27% (OR 1.04; 95% CI 0.76, 1.43; p = 0.79). (Helton TJ, Bavry AA, Kumbhani DJ, et al. Incremental effect of clopidogrel on important outcomes in patients with cardiovascular disease: a meta-analysis of randomized trials. Am J Cardiovasc Drugs. 2007;7:289 –97.)
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Current Submission: An updated literature search covering January 1, 2016 through February 19, 2020 was performed. A search using the MeSH search terms “coronary artery disease” and “acetylsalicylic acid” resulted in 745 articles. A search using the MeSH search terms “coronary artery disease” and “clopidogrel” resulted in 465 articles. A search using the MeSH search terms “coronary artery disease” and “antiplatelet therapy” resulted in 863 articles. None of the studies contained new conclusions that would alter the recommendation to prescribe antiplatelet therapy to patients with coronary artery disease. Last Submission: One meta-analysis was published after the publication of the 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease. Note: Text below for description and results is verbatim from the article abstract. Tang XF, Fan JU, Meng J, Jin C, Yan JQ, Yang YJ. Impact of new oral or intravenous P2Y12 inhibitors and clopidogrel on major ischemic and bleeding events in patients with coronary artery disease: a meta-analysis of randomized trials. Atherosclerosis. 2014;233:568-78.

	Description and Results: Twelve randomized, placebo-controlled studies and two subgroup analyses of included studies on ST-segment elevation myocardial infarction (STEMI) were included. The database consisted of 82,784 patients, with 43,875 (53%) on new oral P2Y12 inhibitors and 38909 (47%) on intravenous P2Y12 inhibitors compared with clopidogrel. The primary efficacy endpoint was major adverse cardiac events (MACEs). The primary safety endpoint was thrombolysis in myocardial infarction (TIMI) major bleeding. New oral P2Y12 inhibitors significantly decreased MACEs (odds ratio: 0.85, p<0.0001 for the whole cohort; OR: 0.77, p=0.04 for STEMI) and all-cause death (OR: 0.88, p=0.04 for the whole cohort; OR: 0.77, p=0.01 for STEMI). Among new intravenous P2Y12 inhibitors, only cangrelor significantly decreased the risk of MACEs. An increase in TIMI major bleeding was observed only by prasugrel among the new P2Y12 inhibitors. Conclusion: New oral P2Y12 inhibitors reduce ischemic events, but there is no obvious increase in major bleeding in patients with CAD, and the risk/benefit ratio is particularly favorable for STEMI patients. Moreover, only cangrelor is beneficial for ischemic events in patients on new intravenous P2Y12 inhibitors. Impact on conclusions of systematic review: This additional meta-analysis does not impact the current guideline recommendations on which this measure is based.

Citations:

1. CAPRIE Steering Committee (1996). A randomised, blinded, trial of clopidogrel versus aspirin in patients at risk of ischaemic events (CAPRIE). CAPRIE Steering Committee. *Lancet (London, England)*, *348*(9038), 1329–1339. https://doi.org/10.1016/s0140-6736(96)09457-3
2. Yusuf, S., Zhao, F., Mehta, S. R., Chrolavicius, S., Tognoni, G., Fox, K. K., & Clopidogrel in Unstable Angina to Prevent Recurrent Events Trial Investigators (2001). Effects of clopidogrel in addition to aspirin in patients with acute coronary syndromes without ST-segment elevation. *The New England journal of medicine*, *345*(7), 494–502. https://doi.org/10.1056/NEJMoa010746

1a.4 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.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

1a.4.2 What process was used to identify the evidence?

1a.4.3. Provide the citation(s) for the evidence.

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., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

In the absence of contraindications, antiplatelet therapy is recommended for patients with a diagnosis of coronary artery disease as it reduces the risk of adverse coronary events, including death, by inhibiting platelet aggregation. Aspirin and clopidogrel have been proven to stabilize coronary plaque and prevent ruptures and clots. Despite the strong evidence, adherence to these guideline-directed therapies is suboptimal.

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for maintenance of endorsement. 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 include.) This information also will be used to address the sub-criterion on improvement (4b1) under Usability and Use.*

A total of 1,846 providers are reporting on this measure through the registry reporting option for MIPS during the period between 1/1/2018-12/31/2018. This data set reflects information at the provider level and our analysis of the data as a whole is reflected throughout this submission. Of those 1,846 providers, 1,843 providers had at least one patient who qualified for the measure, after accounting for exceptions, for a total of 506,259 eligible patients. The average number of eligible patients is 274 for the 1,843 providers. The range of eligible patients for 1,843 providers is from 1 to 2,781.

Based on the sample of 1,843 included providers, the mean performance rate is 0.88, the median performance rate is 0.88 and the mode is 1.0. 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 0.15 (1.00–0.85). Deciles are provided in the table below:

Decile	Performance
1	0.68
2	0.82
3	0.87
4	0.91
5	0.94
6	0.96
7	0.99

8 1.00

9 1.00

10 1.00

In addition, CMS published its quality benchmarks for MIPS 2020, 2019, and 2018. CMS describes that benchmarks are created using historical performance rates. For example, 2020 benchmarks are based on actual performance data submitted to the Quality Payment Program two years prior (i.e., 2018). Note that the performance scores listed in this section are not consistently derived from a nationally representative sample. The average performance rates and standard deviations for this measure from 2016 through 2018 are:

Year	Submission Method	Average Performance Rate	Standard Deviation
2018	CQM	89.2	Not available
2017	Registry/QCDR	89.6	13.2
2016	Registry/QCDR	87.3	11.72013 performance data from the Pinnacle registry.

Overall mean performance on this measure is 86.2%, with a standard deviation of 10.5%. The minimum score equals 0.00%, while the maximum score equals 100.00%. The interquartile score is equal to 10.3%.

2,407 providers were measured, and the patient study sample equals 1,023,530. 62.4% of the sample is male. 92.0% of the sample is white, 5.8% is black, and 2.2 % identified as “other.” The sample reached across all US regions, with 12.7% of providers in the Northeast, 29.0% of providers in the Midwest, 39.7 % of providers in the South, and 18.6% of providers in the West.

Mean

Decile 162.6%

Decile 277.7%

Decile 382.6%

Decile 485.5%

Decile 587.6%

Decile 689.5%

Decile 791.2%

Decile 893.0%

Decile 994.9%

Decile 10 97.6%

2014 performance data from the Pinnacle registry.

Overall mean performance on this measure is 86.3%, with a standard deviation of 9.49%. The minimum score equals 0.00%, while the maximum score equals 100.00%. The interquartile score is equal to 10.2%.

2,248 providers were measured, and the patient study sample equals 959,792. 62.6% of the sample is male. 92.9% of the sample is white, 5.1% is black, and 2.0% identified as “other.” The sample reached across all US regions, with 11.4% of providers in the Northeast, 28.5% of providers in the Midwest, 40.3 % of providers in the South, and 19.8% of providers in the West.

Mean

Decile 165.4%

Decile 277.8%

Decile 382.4%

Decile 485.1%

Decile 587.3%

Decile 689.2%

Decile 791.0%

Decile 892.8%

Decile 994.7%

Decile 10 97.4%

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.

A study that looked at adherence to cardiovascular performance measures within the PINNACLE Registry found gaps in care in prescribing patterns amongst clinicians. The study included 1,493 individual practitioners who saw 769,139 patients; among patients with coronary artery disease (CAD), 72% treated by male clinicians received antiplatelet therapy and 66% of eligible patients treated by female clinicians received antiplatelet therapy (P<0.001) (Gupta et al., 2018), indicating not only gaps in care but disparities in prescribing patterns amongst clinicians. A similar study analyzed adherence to clinical practice guidelines using data from the PINNACLE Registry and found that among 525,761 eligible patients with CAD, 77% of patients received antiplatelet therapy (Fleming et al., 2016). Tran et al. evaluated trends in cardiac medication adherence amongst 1,789 American adults with a history of CAD. Using data from the National Health and Nutrition Examination Survey, authors found that between 2011 and 2014, 69% (SE = 2.4) of American adults with CAD reported using aspirin (2017). Lastly, other researchers evaluated the utilization of antiplatelet medications among over 10,000 adults with coronary heart disease, using national data from the Medical Expenditure Panel Survey between 2003 to 2012. 71% of respondents reported using antiplatelet therapy (95% CI 69.7 to 72.1), 68% reported using aspirin (95% CI 66.3 to 68.8) and 17% (95% CI 15.8 to 17.7) reported using either clopidogrel, ticlopidine, or prasugrel (Johansen, Hefner, & Foraker, 2015).

Citations:

1. Gupta, D., Tang, F., Masoudi, F. A., Jones, P. G., Chan, P. S., & Daugherty, S. L. (2018). Practitioner Gender and Quality of Care in Ambulatory Cardiology Practices: A Report From the National Cardiovascular Data Practice Innovation and Clinical Excellence (PINNACLE) Registry. The Journal of Cardiovascular Nursing, 33(3), 255–260. https://doi.org/10.1097/JCN.0000000000000443

2. Fleming, L. M., Jones, P., Chan, P. S., Andrei, A. C., Maddox, T. M., & Farmer, S. A. (2016). 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. Circulation. Cardiovascular Quality and Outcomes, 9(1), 48–54. https://doi.org/10.1161/CIRCOUTCOMES.115.002159

3. Johansen, M. E., Hefner, J. L., & Foraker, R. E. (2015). Antiplatelet and Statin Use in US Patients With Coronary Artery Disease Categorized by Race/Ethnicity and Gender, 2003 to 2012. The American Journal of Cardiology, 115(11), 1507–1512. https://doi.org/10.1016/j.amjcard.2015.02.052

4. Tran, H. V., Waring, M. E., McManus, D. D., Erskine, N., Do, V., Kiefe, C. I., & Goldberg, R. J. (2017). Underuse of Effective Cardiac Medications Among Women, Middle-Aged Adults, and Racial/Ethnic Minorities With Coronary Artery Disease (from the National Health and Nutrition Examination Survey 2005

to 2014). The American journal of cardiology, 120(8), 1223–1229. https://doi.org/10.1016/j.amjcard.2017.07.004

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 maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b1) 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. See section 1b.5. for disparities data from the literature.

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in 1b.4

The Tran et al. study mentioned previously, which analyzed data from the National Health and Nutrition Examination Survey, found that among those with coronary artery disease (CAD), women and racial/ethnic minorities were less likely to take aspirin, compared with men and non-Hispanic Whites (OR = 0.63 and 0.74 respectively) (2017).

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 sub criteria 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 : Coronary Artery Disease

De.6. Non-Condition Specific*(check all the areas that apply):*

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

Adults, Elderly

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 with this submission. Additional measure details may be found at https://qpp.cms.gov/mips/explore-measures/quality-measures

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.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:**

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

S.3.1. <u>For maintenance of endorsement:</u> Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

Yes

S.3.2. <u>For maintenance of endorsement,</u> please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

Supporting guidelines and coding included in the measure are reviewed on an annual basis. This annual review has resulted in minor changes to the coding.

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) DO NOT include the rationale for the measure.*

<u>IF an OUTCOME MEASURE</u>, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Patients who were prescribed aspirin or clopidogrel

S.5. 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, time period for data collection, 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)*

<u>IF an OUTCOME MEASURE,</u> describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Time period for data collection: At least once during the measurement period

Definition: Prescribed - May include prescription given to the patient for aspirin or clopidogrel at one or more visits in the measurement period OR patient already taking aspirin or clopidogrel as documented in current medication list. Report CPT Category II code 4086F: Aspirin or clopidogrel prescribed

S.6. 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

S.7. Denominator Details *(All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, 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.)*

<u>IF an OUTCOME MEASURE</u>, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Time period for data collection: 12 consecutive months

Patients aged >/= 18 years on date of encounter

AND

Diagnosis for coronary artery disease (ICD-10-CM): 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, I21.9, I21.A9, 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 performance 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

WITHOUT

Telehealth Modifier: GQ, GT, 95, POS 02

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

Denominator exceptions

- Documentation of medical reason(s) for not prescribing aspirin or clopidogrel (e.g., allergy, intolerance, receiving other thienopyridine therapy, receiving warfarin therapy, bleeding coagulation disorders, other medical reasons)
- Documentation of patient reason(s) for not prescribing aspirin or clopidogrel (e.g., patient declined, other patient reasons)
- Documentation of system reason(s) for not prescribing aspirin or clopidogrel (e.g., lack of drug availability, other reasons attributable to the health care system)

S.9. Denominator Exclusion Details *(All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, 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.)*

Time period for data collection: During the encounter within the 12-month period

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 AHA and ACC 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 Coronary Artery Disease (CAD): Antiplatelet Therapy, exceptions may include medical reason(s) (e.g., allergy, intolerance, receiving other thienopyridine therapy, receiving warfarin therapy, bleeding coagulation disorders, other medical reasons), patient reason(s) (e.g., patient declined, other patient reasons), or system reason(s) (e.g., lack of drug availability, other reasons attributable to the health care system) for not prescribing aspirin or clopidogrel. Although this methodology does not require the external

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

Additional details:

Append a modifier to CPT Category II code:

4086F-1P: Documentation of medical reason(s) for not prescribing aspirin or clopidogrel (e.g., allergy, intolerance, receiving other thienopyridine therapy, receiving warfarin therapy, bleeding coagulation disorders, other medical reasons)

4086F-2P: Documentation of patient reason(s) for not prescribing aspirin or clopidogrel (e.g., patient declined, other patient reasons)

4086F-3P: Documentation of system reason(s) for not prescribing aspirin or clopidogrel (e.g., lack of drug availability, other reasons attributable to the health care system)

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate – 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 national recommendations put forth by the IOM (now NASEM) 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.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Rate/proportion

If other:

S.13. 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.14. Calculation Algorithm/Measure Logic (*Diagram or 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; time period for data, aggregating data; risk adjustment; etc.)*

Calculating the performance rate:

1. Define the initial population. The initial population is identified through a common set of characteristics that define the overall group of patients – or other unit of measurement – targeted for evaluation

2. Define the denominator by identifying the subset of the initial population that meets the denominator criteria. Note: in some cases, the initial population and denominator are identical

3. Determine the numerator by identifying the subset of the denominator that meets the numerator criteria

4. From the patients who did not meet the numerator criteria, determine if the provider has documented whether each patient represents an exception. Subtract from the denominator those patients that meet the conditions for a denominator exception; although the exception cases are removed from the denominator for the measure calculation, the exception rate (i.e., percentage of patients with valid exceptions) should be calculated and reported along with performance rates to highlight variations in care

5. Calculate the performance rate

A patient not meeting the numerator criteria and without a valid and documented exception represents a quality failure.

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

Not applicable. This measure is not based on a sample

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

Not Applicable. This measure is not based on a survey.

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Registry Data

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

Quality Payment Program – MIPS Quality Program.

S.19. 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.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Clinician : Individual

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Home Care, Outpatient Services, Post-Acute Care

If other:

S.22. 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

2. Validity – See attached Measure Testing Submission Form

CAD_0067_Testing_Attachment_2020_Final_revisions.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

No - This measure is not risk-adjusted

Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 0067
Measure Title: Coronary Artery Disease (CAD): Antiplatelet Therapy
Date of Submission: 1/6/2020
Type of Measure:

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

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.17***)	**Measure Tested with Data From:**

☐ abstracted from paper record	☐ abstracted from paper record
☐ claims	☐ claims
☒ registry	☒ 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*).

Previous testing

The primary analysis was conducted at the level of the individual provider and included all patients with coronary artery disease (CAD) cared for by that provider and captured in the PINNACLE Registry during the one-year study period. The PINNACLE Registry systematically maps each practice's Electronic Health Record to the data elements required for the 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 and are available for CMS reporting. Using these data, we were able to calculate the number of patients who should have received antiplatelet therapy, or a clinically, evidence-based reason not to use antiplatelet therapy was documented. This means that every 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 in these analyses and meeting the performance measure on any single visit within the year met the criterion for this measure.

Current testing

The data source is Registry data from the 2018 Merit-based Incentive Payment System (MIPS) Program. This data source was the most comprehensive source available at the time of analysis. Our analysis was limited to include only data that was reported at the unique NPI level.

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

Previous 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.

Current testing

The data are for the time period January 1st, 2018 through December 31st, 2018.

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.20*)	**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*)

Previous testing

2013

2,407 providers met the minimum number of eligible patients (10) for inclusion in the reliability analysis. The average number of eligible patients for providers included is 425.2 for a total of 1,023,530 patients. The range of number of patients for providers included is from 10 to 2,834. As described above, providers with fewer than 10 eligible patient encounters during the study period were excluded.

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 = 2407
Provider gender	
(1) Male	1923 (80.0%)
(2) Female	482 (20.0%)
Missing (.)	2
Provider	
NP/PA	258 (10.9%)
MD/DO	2061 (86.9%)
RN/nurses	52 (2.2%)
Missing (.)	36
Region	
(1) Northeast	305 (12.7%)
(2) Midwest	698 (29.0%)
(3) South	956 (39.7%)

(4) West	448 (18.6%)

2014

2,248 providers met the minimum number of eligible patients (10) for inclusion in the reliability analysis. The average number of eligible patients for providers included is 427.0 for a total of 959,792 patients.

The range of numbers of patients for providers included is from 10 to 2,649. As described above, providers with fewer than 10 eligible patient encounters during the study period were excluded.

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 = 2248
Provider gender	
(1) Male	1784 (79.4%)
(2) Female	464 (20.6%)
Provider	
NP/PA	250 (11.3%)
MD/DO	1915 (86.7%)
RN/nurses	44 (2.0%)
Missing (.)	39
Region	
(1) Northeast	257 (11.4%)
(2) Midwest	640 (28.5%)
(3) South	905 (40.3%)
(4) West	446 (19.8%)

Current testing

A total of 1,846 providers are reporting on this measure through the registry reporting option for MIPS during the period between 1/1/2018-12/31/2018. This data set reflects information at the provider level and our analysis of the data as a whole is reflected throughout this submission. Of those 1,846 providers, 1,843 providers had at least one patient who qualified for the measure, after accounting for exceptions, for a total of 506,259 eligible patients. The average number of eligible patients is 274 for the 1,843 providers. The range of eligible patients for 1,843 providers is from 1 to 2,781.

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*)

Previous testing

2013

There are a total of 1,023,530 patients included in the temporal comparison that were treated in 2013. Patients' characteristics are provided below:

	Total
	n = 1023530
Race	
(1) White	502521 (92.0%)
(2) Black	31466 (5.8%)
(3) Other	12014 (2.2%)
Missing (.)	170072
Insurance	
(0) No insurance	1381 (1.0%)
(1) Private	100517 (69.9%)
(2) Medicare	40963 (28.5%)
(3) Medicaid	182 (0.1%)
(4) Other	820 (0.6%)
Missing (.)	154582
Age	
18 to <60	197507 (19.3%)
60 to <70	285197 (27.9%)
70 to <80	320245 (31.3%)
80 to 114	220581 (21.6%)
Sex	
(1) Male	637419 (62.4%)
(2) Female	384062 (37.6%)
Missing (.)	1918
BMI (kg/m2)	30.6 ± 9.0
Missing	231647
Diabetes Mellitus)	288906 (28.9%)
Hypertension	807299 (85.4%)
Atrial Fibrillation/Flutter	217128 (22.7%)
Heart Failure	239885 (24.5%)
Peripheral Arterial Disease	139911 (16.1%)
	Total
	n = 1023530

Stroke/TIA	40229 (6.1%)
Myocardial Infarction	261043 (30.3%)

2014

There are a total of 959,792 patients included in the primary analysis (2014). See below for details on patient characteristics.

	Total **n = 959792**
Race	
(1) White	428912 (92.9%)
(2) Black	23387 (5.1%)
(3) Other	9331 (2.0%)
Missing (.)	158682
Insurance	
(0) No insurance	528 (0.4%)
(1) Private	99826 (81.6%)
(2) Medicare	21467 (17.6%)
(3) Medicaid	177 (0.1%)
(4) Other	284 (0.2%)
Missing (.)	158102
Age	
18 to <60	179194 (18.7%)
60 to <70	265376 (27.6%)
70 to <80	306202 (31.9%)
80 to 114	209020 (21.8%)
Sex	
(1) Male	599619 (62.6%)
(2) Female	357647 (37.4%)
Missing (.)	2248
BMI (kg/m2)	30.7 ± 8.9
Missing	189212
Diabetes Mellitus	265083 (29.4%)
Hypertension	724628 (86.1%)

	Total n = 959792
Atrial Fibrillation/Flutter	210886 (23.9%)
Heart Failure	233017 (26.1%)
Peripheral Arterial Disease	136158 (16.3%)
Stroke/TIA	40536 (6.2%)
Myocardial Infarction	238396 (29.3%)

Current testing

There were 506,259 patients included in this reliability testing and analysis. These were the patients that were associated with providers who had at least one eligible patient in the year.

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.

Previous testing

The dataset described above was used for all aspects of testing.

Current testing

The same data samples were used for reliability testing and exceptions analysis.

1.8 What were the social risk factors that were available and analyzed? For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

Previous testing

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.

Current testing

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*)

Previous testing

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], where the latter represents the within-physician estimate of our error in assessing their

'true' performance. Thus, the reliability estimate 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 five different distributions of providers' patient volumes: 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.

Current testing

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 provider performance and the noise is the total variability in measured performance. Reliability at the level of the specific provider is given by:

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

Reliability is the ratio of the provider-to-provider variance divided by the sum of the provider-to-provider variance plus the error variance specific to a provider. 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 provider performance.

Variance (provider-to-provider) = alpha*beta/ ((alpha + beta + 1) * (alpha + beta) ^2)

Variance (provider-specific-error) = p(1-p)/n

Where p is the passing rate for a provider and n is the number of patients for that provider

Reliability testing was performed by using a beta-binomial model. The beta-binomial model assumes the provider performance score is a binomial random variable conditional on the provider'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.

For this analysis Alpha = 2.7189 and Beta = 0.4192. These parameters are used to calculate the variance (provider-specific-error) which is approximately equal to 0.002. Reliability is then calculated for each provider using this value and the variance (provider-to-provider). Average reliability is reported by averaging reliability for each provider with at least 1 patient for the measure.

A reliability of 0.70 – 0.80 is generally considered the acceptable threshold for reliability, 0.80 – 0.90 is considered high reliability, and 0.90 – 1.0 is considered very high. [1]

1. Adams JL, Mehrotra A, McGlynn EA, Estimating Reliability and Misclassification in Physician Profiling, Santa Monica, CA: RAND Corporation, 2010. www.rand.org/pubs/technical_reports/TR863. (Accessed on February 24, 2012.)

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*)

Previous testing

2013 – In 2013, the signal-noise ratios are shown below:

Description	Number of Patients	Signal-to-Noise Ratio
Minimum	10	0.994
25th percentile	188	0.998
50th percentile	368	0.998
75th percentile	586	0.999
Average	426	0.998

2014 – In 2014, the signal-noise ratios are shown below:

Description	Number of Patients	Signal-to-Noise Ratio
Minimum	10	0.995
25th percentile	204	0.998
50th percentile	376	0.998
75th percentile	578	0.999
Average	427	0.998

Current testing

The average reliability for providers with at least one eligible patient is 0.95. We also report the average reliability at each decile of the sample shown in the table below.

Reliability Statistics														
Denominator (Patients)	Mean	SD	Min	Max	Decile									
					1	2	3	4	5	6	7	8	9	10
1+	0.96	0.11	0.18	1.00	0.70	0.93	0.97	0.99	0.99	1.00	1.00	1.00	1.00	1.00

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?*)

Previous testing

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 patient visits required (>10) the average reliability was 0.994 and 0.986 for 2013 and 2014, respectively. For providers with the median number of patient encounters, the reliability was even higher, with 0.998 for both years. 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 true differences across physicians.

Current testing

This measure has very high reliability and increases as the denominator size increases for both years.

2b1. VALIDITY TESTING

2b1.1. 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*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. 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)*

Previous testing

Content validity for this measure was assessed by expert work group members during the development process. 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 as needed. 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 and voting by the PCPI membership.

Members of the expert work group that developed the measure included: Bruce Abramowitz, MD, FACC (interventional cardiology; measure implementation), Karen Alexander, MD (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 (guideline development), Michael O'Toole, MD (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).

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 addition 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. In 2013 the rate of records not passing the quality evaluation was 1.9% and in 2014 it was 2.1%. Feedback reports are generated to facilitate investigation of data issues at the practice level to verify accuracy of abstraction and to remap elements whose definitions or recording have changed.

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

Forty Two (42) members completed the survey and provided a mean importance rating of 4.26, with 83.3% agreeing with the use of the measure for quality assessment.

Current testing

Validity testing method

Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) (NQF 0066) was chosen as a suitable candidate for correlation analysis due to the similarities in patient population and domain. We hypothesize that there exists a positive association of scores between providers who prescribe an antiplatelet therapy on patients with coronary artery disease and those who prescribe Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy on patients with diabetes or left ventricular systolic dysfunction (LVEF < 40%) within a 12 month period.

Providers included in the analysis had at least one patient in the denominator after exceptions were removed. Datasets were reviewed to identify providers based on the provider identifier. Comparing performance scores of those shared provider IDs, the empirical analysis uses regression with Coronary Artery Disease (CAD): Antiplatelet Therapy (NQF 0067) as the outcome and Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) (NQF 0066) as the predictor. Results identify the multiple R value (the correlation coefficient) and P-value of the regression variables to assess the association between performance scores of these shared provider IDs.

We use the following guidance to describe correlation[1]:

Correlation	Interpretation
0.80 – 1.00	Very Strong
0.60 – 0.79	Strong
0.40 – 0.59	Moderate
0.20 - 0.39	Weak
0 – 0.19	Very Weak

1. "11. Correlation and Regression." *The BMJ*, 21 March 2019, https://www.bmj.com/about-bmj/resources-readers/publications/statistics-square-one/11-correlation-and-regression/ .

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

Previous testing

We believe that the processes used to extract data from the exact source from which any abstraction process done manually would use (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 = 42; Mean rating = 4.26 and 83.3% of respondents either agree or strongly agree that this measure can accurately distinguish good and poor quality.

Frequency Distribution of Ratings 1 - <2> (Strongly Disagree)

2 - <3>

3 - <2> (Neither Agree nor Disagree) 4 - <10>

5 - <25> (Strongly Agree)

Current Testing

Data from the 2018 MIPS Registry Program were used to perform the correlation analysis for this measure. Data comes from the Registry version of Coronary Artery Disease (CAD): Antiplatelet Therapy (NQF 0067) and Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) (NQF 0066).

Coronary Artery Disease (CAD): Antiplatelet Therapy (NQF 0067) was positively correlated with Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) (NQF 0066).

NQF #0067

Coefficient of correlation = 0.47
Alpha level = 0.05
P-value = < 0.001
Number of shared Providers based on Provider identifier = 520

2b1.4. 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?*)

Previous testing

The measure was judged to have high face validity by both its clinical importance and by the 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.

Current Testing

Coronary Artery Disease (CAD): Antiplatelet Therapy (NQF 0067) has a moderate positive correlation with Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) (NQF 0066). The correlation is highly statistically significant. With a coefficient of correlation of 0.47, the correlation is moderate, significant, and confirms our hypothesis. The moderate positive correlation with Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) (NQF 0066) demonstrates the criterion validity of the measure. The strength of the correlation is within our expectations.

2b2. EXCLUSIONS ANALYSIS

NA ☐ no exclusions — *skip to section 2b3*

2b2.1. 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*)

Previous testing

Since not all patients with CAD will meet the guideline recommendations for antiplatelet therapy, exclusions in this measure are intended to remove patients for whom antiplatelet therapy may not be appropriate. We divide these into two categories: Exclusions and Exceptions. Exclusions arise when patients who are included in the initial patient or eligible population for the measure set do not meet the denominator criteria specific to the intervention required by the numerator. Exclusions are absolute and apply to all patients and therefore are not part of clinical judgment within a measure. Specific exclusions should be derived from evidence-based guidelines. Exceptions are not absolute, and are based on clinical judgment and individual patient characteristics, thus patients with such contraindications represent circumstances where the clinicians balanced the risks and benefits and felt that, in a given situation, the benefits outweighed the risks and chose

to treat the patient. These patients are therefore included in both the numerator and denominator of the measure. In contrast, the exceptions are clearly documented reasons to not treat the patient and are removed from the denominator of the population.

Exclusions in this measure:

- Documented medical, patient, or system reason for not on aspirin or clopidogrel or patients were on warfarin.

Current Testing

This measure does not include exclusions and does specify exceptions that we analyzed.

Exceptions include:

- Documentation of medical reason(s) for not prescribing aspirin or clopidogrel (e.g., allergy, intolerance, receiving other thienopyridine therapy, receiving warfarin therapy, bleeding coagulation disorders, other medical reasons)
- Documentation of patient reason(s) for not prescribing aspirin or clopidogrel (e.g., patient declined, other patient reasons)
- Documentation of system reason(s) for not prescribing aspirin or clopidogrel (e.g., lack of drug availability, other reasons attributable to the health care system)

Exceptions were analyzed for frequency across providers and deciles of exceptions were reported.

2b2.2. 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*)

Previous testing

The Exceptions for each year are provided below:

2013: 3.1% (n=74) of the providers do not have exceptions in the denominator. Among 2,333 providers who do have exceptions, the exception rate ranges from 0.2% to 35.3%, mean is 6.0%. Of those patients removed from the measure, 0.72% were removed due to a medical reason, 9.47% were removed for a patient reason, 1.48% were removed for a system reason, 0.30% were removed due to multiple reasons, and 88.03% were removed due to patients who were on warfarin or another thienopyridine that is not included in this measure (e.g. prasugrel or ticagrelor) and the physician felt that the addition of aspirin or a clopidogrel provided an increased risk of bleeding with minimal benefits in the prevention of recurrent cardiovascular events.

2014: 2.5% (n=56) of the providers do not have exceptions in the denominator. Among 2,192 providers who do have exceptions, the exception rate ranges from 0.3% to 27.4%, mean is 5.8%. Of those patients removed from the measure, 0.80% were removed due to a medical reason, 9.52% were removed for a patient reason, 1.88% were removed for a system reason, 0.31% were removed due to multiple reasons, and 87.50% were removed due to concomitant medications that either provide anti-platelet or systemic anti-coagulation.

Current Testing

Amongst the 1,843 included providers, there were a total of 564 exceptions reported. The average number of exceptions per provider in this sample is 11.22. The proportion of exceptions to patients is 0.04. Exception deciles illustrate the spread of exceptions amongst providers. According to the results, 50% of providers had 3 or fewer exceptions across eligible patients for the year under study.

Decile	Exceptions
1	0
2	0
3	0
4	0
5	1
6	4
7	8
8	15
9	30
10	564

2b2.3. 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*)

Previous testing

We do not view any concerns with exclusions or exceptions for this measure. While the majority of exceptions are due to patient reasons and these might be 'gameable' by clinicians, we recognize the bleeding and bruising are very frequent side effects of anti-platelet treatment (see Amin et al, J Am Coll

Cardiol 2013; 2013; 61(21):2130-8 and that clinicians who think enough about this decision process to document an exclusion are likely providing excellent, patient-centered care.

Current Testing

The AHA follows the PCPI methodology in distinguishing between denominator exceptions and denominator exclusions.

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 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 Coronary Artery Disease (CAD): Antiplatelet Therapy, exceptions may include medical reason(s) (e.g., allergy, intolerance, receiving other thienopyridine therapy, receiving warfarin therapy, bleeding coagulation disorders, other medical reasons), patient reason(s) (e.g., patient declined, other patient reasons), or system reason(s) (e.g., lack of drug

availability, other reasons attributable to the health care system) for not prescribing aspirin or clopidogrel. 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.

2b3. 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 2b4.

2b3.1. 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

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

Previous testing

Not applicable

Current Testing

Not applicable

2b3.2. 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.

Previous testing

Not applicable

Current Testing

Not applicable

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk 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*) **Also discuss any "ordering" of risk factor inclusion**; for example, are social risk factors added after all clinical factors?

Previous testing

Not applicable

Current Testing

Not applicable

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

- ☐ **Published literature**
- ☐ **Internal data analysis**
- ☐ **Other (please describe)**

Current Testing

Not applicable

2b3.4a. What were the statistical results of the analyses used to select risk factors?

Previous testing

Not applicable

Current Testing

Not applicable

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk 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.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

Previous testing

Not applicable

Current Testing

Not applicable

2b3.5. 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*)

Previous testing

Not applicable

Current Testing

Not applicable

Provide the statistical results from testing the approach to controlling for differences in patient characteristics (case mix) below.

If stratified, skip to 2b3.9

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

Previous testing

Not applicable

Current Testing

Not applicable

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

Previous testing

Not applicable

Current Testing

Not applicable

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

Previous testing

Not applicable

Current Testing

Not applicable

2b3.9. Results of Risk Stratification Analysis:

Previous testing

Not applicable

Current Testing

Not applicable

2b3.10. 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*)

Previous testing

Not applicable

Current Testing

Not applicable

2b3.11. 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*)

Previous testing

Not applicable

Current Testing

Not applicable

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. 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)*

Previous testing

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	2403	637419	0.00%	85.7%	88.6%	94.7%	100%	9.02%	9.97%
Female	2403	384062	0.00%	77.1%	82.5%	90.8%	100%	13.8%	12.4%
Age: <60	2404	197507	0.00%	73.9%	81.6%	92.7%	100%	18.7%	15.1%
Age: 60 -< 70	2406	285197	0.00%	83.9%	87.5%	94.5%	100%	10.6%	11.1%
Age: 70 -< 80	2405	320245	0.00%	84.6%	87.9%	94.4%	100%	9.80%	10.6%
Age: >= 80	2400	220581	0.00%	83.0%	87.1%	94.4%	100%	11.4%	11.3%
Insurance: None	134	1381	0.00%	92.3%	86.1%	100%	100%	7.69%	28.6%
Insurance: Private	683	100517	0.00%	83.9%	87.5%	96.0%	100%	12.1%	13.8%
Insurance: Medicaid	500	40963	0.00%	84.7%	89.5%	100%	100%	15.3%	13.1%
Insurance: Medicare	27	182	50.0%	86.7%	92.3%	100%	100%	13.3%	12.5%
Insurance: Other	152	820	0.00%	68.3%	81.4%	100%	100%	31.7%	29.2%
Race: White	1552	502521	0.00%	83.7%	86.8%	93.7%	100%	10.1%	10.9%
Race: Black	1425	31466	0.00%	77.8%	83.9%	100%	100%	22.2%	19.9%

Race: Other	1190	12014	0.00%	75.0%	83.8%	100%	100%	25.0%	25.6%

2014

Label	# of providers	# of patients	Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std Dev
Male	2243	599619	0.00%	85.7%	88.8%	94.5%	100%	8.80%	8.84%
Female	2242	357647	0.00%	76.6%	82.2%	90.6%	100%	14.0%	11.7%
Age: <60	2246	179194	0.00%	74.7%	82.0%	92.1%	100%	17.3%	14.2%
Age: 60 - < 70	2248	265376	0.00%	83.7%	87.5%	94.1%	100%	10.5%	10.0%
Age: 70 - < 80	2248	306202	20.0%	83.9%	87.9%	94.3%	100%	10.4%	9.50%
Age: >= 80	2242	209020	0.00%	83.0%	87.1%	93.9%	100%	10.9%	10.5%
Insurance: None	104	528	0.00%	100%	90.5%	100%	100%	0.00%	24.1%
Insurance: Private	606	99826	0.00%	84.3%	88.6%	96.5%	100%	12.1%	12.2%
Insurance: Medicaid	397	21467	0.00%	85.0%	89.8%	100%	100%	15.0%	13.4%
Insurance: Medicare	27	177	0.00%	80.0%	87.9%	100%	100%	20.0%	21.5%
Insurance: Other	103	284	0.00%	77.8%	83.0%	100%	100%	22.2%	31.8%
Race: White	1288	428912	0.00%	83.7%	87.1%	93.3%	100%	9.58%	9.79%
Race: Black	1212	23387	0.00%	77.2%	84.7%	100%	100%	22.8%	18.4%
Race: Other	1090	9331	0.00%	76.9%	84.6%	100%	100%	23.1%	25.4%

Current Testing

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

2b4.2. 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*)

2013

Overall mean performance on this measure is 86.2%, with a standard deviation of 10.5%. The minimum score equals 0.00%, while the maximum score equals 100.00%. The interquartile score is equal to 10.3%. 2,407 providers were measured, and the patient study sample equals 1,023,530. 62.4% of the sample is male. 92.0% of the sample is white, 5.8% is black, and 2.2 % identified as "other." The sample reached across all US regions, with 12.7% of providers in the Northeast, 29.0% of providers in the Midwest, 39.7% of providers in the South, and 18.6% of providers in the West.

# of providers	Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std Dev
2407	0.00%	82.6%	86.2%	92.9%	100%	10.3%	10.5%

	Mean
Decile 1	62.6%
Decile 2	77.7%
Decile 3	82.6%
Decile 4	85.5%
Decile 5	87.6%
Decile 6	89.5%
Decile 7	91.2%
Decile 8	93.0%
Decile 9	94.9%
Decile 10	97.6%

2014

Overall mean performance on this measure is 86.3%, with a standard deviation of 9.49%. The minimum score equals 0.00%, while the maximum score equals 100.00%. The interquartile score is equal to 10.2%.

2,248 providers were measured, and the patient study sample equals 959,792. 62.6% of the sample is male. 92.9% of the sample is white, 5.1% is black, and 2.0% identified as "other." The sample reached across all US regions, with 11.4% of providers in the Northeast, 28.5% of providers in the Midwest, 40.3% of providers in the South, and 19.8% of providers in the West.

# of providers	Minimum	Lower Quartile	Mean	Upper Quartile	Maximum	Quartile Range	Std Dev
2248	10.0%	82.6%	86.3%	92.8%	100%	10.2%	9.49%

	Mean
Decile 1	65.4%
Decile 2	77.8%
Decile 3	82.4%
Decile 4	85.1%
Decile 5	87.3%
Decile 6	89.2%
Decile 7	91.0%
Decile 8	92.8%
Decile 9	94.7%
Decile 10	97.4%

Current Testing

Based on the sample of 1,843 included providers, the mean performance rate is 0.88, the median performance rate is 0.88 and the mode is 1.0. 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 0.15 (1.00–0.85). Deciles are provided in the table below:

Decile	Performance
1	0.68
2	0.82
3	0.87
4	0.91
5	0.94
6	0.96
7	0.99
8	1.00
9	1.00
10	1.00

2b4.3. 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?*)

Previous testing

2013: A small amount of variability was noted among providers. The performance-met rate range was 0- 100% with the inter-quartile range being 82.6% to 92.9%. This yielded a Median Rate Ratio of 1.11 (1.10, 1.11). The Median Rate Ratio measures the variation across providers for statistically 'identical' patients and suggests that a patient presenting to 1 provider, as opposed to another, would, on average, be 11% more likely to be treated with an anti-platelet agent.

2014: A small amount of variability was also noted among providers in 2014. The performance-met rate range was 10-100% with the inter-quartile range being 83% to 93%. This yielded a Median Rate Ratio of 1.09 (1.09, 1.10).

Current Testing

The range of performance from 0.00 to 1.00 suggests there's clinically meaningful variation across providers' performance. Outliers are considered to be values less than quartile 1 (0.85) or greater than quartile 3 (1.00) by more than 1.5 the IQR (0.15). Deciles are provided in the table below:

Quartile	Performance
1	0.85
2	0.94
3	1.00
4	1.00

2b5. 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 social risk 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 social risk 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.***

2b5.1. 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*)

Previous testing

Not applicable

Current Testing

Not applicable

2b5.2. 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*)

Previous testing

Not applicable

Current Testing

Not applicable

2b5.3. 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*)

Previous testing

Not applicable

Current Testing

Not applicable

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. 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*)

Previous testing

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 antiplatelet indicates a failure to meet the measure. It is possible that a provider may not have documented antiplatelet therapy use in their EMR system, perhaps if it was provided by another provider in a different healthcare system. However, we believe that this is unlikely and that it is still incumbent upon a provider to clearly document all the medications that a patient is taking, particularly antiplatelet for CAD patients.

Current Testing

The MIPS dataset provided to us by CMS did not contain missing data, so this test was not performed. Nevertheless, missing data may have been rejected when submitted to CMS in which case those values would not be counted towards measure performance. There is no indication that this missing data was systematic, thus their omission would lead to unbiased performance results.

2b6.2. 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*)

Previous testing

Given our assumptions, noted above, we did not conduct an empirical analysis of frequency or distribution of missing data. For this measure, missing data is reported as a quality failure.

Current Testing

This test was not performed for this measure. There was no missing data.

2b6.3. 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*)

Previous testing

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.

Current Testing

The MIPS dataset provided to us by CMS did not contain missing data, so this test was not performed. Nevertheless, missing data may have been rejected when submitted to CMS in which case those values would not be counted towards measure performance. There is no indication that this missing data was systematic, thus their omission would lead to unbiased performance results.

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), 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*) Update this field for **maintenance of endorsement**.

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. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

All the data elements needed for this measure are collected through electronic data.

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. Please also complete and attach the NQF Feasibility Score Card.

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. Required for maintenance of endorsement. Describe difficulties (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 instrument-based, consider implications for both individuals providing 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 owned and copyrighted by the AHA, ACC, and PCPI Foundation, 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 AHA. The AHA 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.

Specific Plan for Use	Current Use (for current use provide URL)
Quality Improvement (Internal to the specific organization)	Public Reporting https://qpp.cms.gov/mips/quality-measures Merit-based Incentive Payment System (MIPS) Payment Program Merit-based Incentive Payment System (MIPS) https://qpp.cms.gov/mips/quality-measures Merit-based Incentive Payment System (MIPS) https://qpp.cms.gov/mips/quality-measures Regulatory and Accreditation Programs Physician Quality Reporting System http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment- Quality Improvement (external benchmarking to organizations) PINNACLE Registry® https://cvquality.acc.org/NCDR-Home/registries/outpatient-registries/pinnacle-registry

4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:

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

- Level of measurement and setting

1. NAME: Merit-based Incentive Payment System (MIPS)-Sponsored by the Centers for Medicare and Medicaid Services (CMS).
PURPOSE: MIPS-Quality. The purpose of the program is to tie payments to clinicians to quality and cost-efficient care, drive improvement in care processes, increase the use of healthcare data, and reduce the costs of care. The program began in 2017.
GEOGRAPHIC AREA: National. Clinicians are included in MIPS-Quality if they are an eligible clinician type and meet program requirements.
LEVEL OF MEASUREMENT: Clinician-level. Professional services rendered under Medicare Part B.
2. NAME: The PINNACLE (Practice Innovation and Clinical Excellence) Registry sponsored by American College of Cardiology and its National Cardiovascular Data Registry.
PURPOSE: The PINNACLE Registry was developed as an outpatient-based prospective quality improvement registry, for cardiology practices in the outpatient setting. Data is collected specifically for coronary artery disease, hypertension, heart failure and atrial fibrillation. Data collected includes patient demographics, medical history, vital signs, laboratory values, imaging results, medications, and contraindications to medications. Registry began in 2008.
GEOGRAPHIC AREA: National. Includes cardiac outpatients. In 2013, the PINNACLE Registry contained information on 2,898,505 patients, cared for by 4,859 providers in 431 practices. By 2017, the registry contained information on 6,040,996 patients, cared for by 8,853 providers in 724 practices.
LEVEL OF MEASUREMENT: Clinician and outpatient care.

4a1.2. 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.

4a1.3. 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.

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

How many and which types of measured entities and/or others were included? If only a sample of measured entities were included, describe the full population and how the sample was selected.

Performance results are provided to all participating clinicians across quality programs. Clinicians participating in the MIPS-Quality program receive performance results, i.e. their MIPS score and corresponding payment adjustment annually. The PINNACLE Registry provides user-friendly online benchmark reports to users via an interactive portal. Individual practice performance on measures is provided by quarter, alongside the national average by quarter. This allows clinicians to identify areas for improvement.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

The feedback process is described in Section 4a2.1.1 and includes providing user-friendly online benchmark reports to users via an interactive portal. Individual practice performance on measures is provided by quarter, alongside the national average by quarter.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

Throughout the year, the American Heart Association receives feedback on its measures from a variety of sources:

1. Ad-hoc tickets/emails sent to the QPP help desk from measure implementers
2. Feedback during public comment periods of regulatory cycles
3. Feedback from the ACC/AHA Task Force on Performance Measures

4a2.2.2. Summarize the feedback obtained from those being measured.

We have received no feedback from other users that resulted in any changes to this measure.

4a2.2.3. Summarize the feedback obtained from other users

We have received no feedback from other users that resulted in any changes to this measure.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

Not applicable based on answers provided in 4a2.2.2 and 4a2.2.3.

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.

4b1. Refer to data provided in 1b but do not repeat here. Discuss any progress on improvement (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.)

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.

The intent of this measure is to improve care amongst individuals with coronary artery disease. The MIPS quality benchmarks cited in 1b2 of this form show a slight improvement in average performance rates between 2016 and 2018. However, reporting rates represent but one facet of the quality improvement process. While we create measures with the goal of improving quality of care, measurement is a mechanism to drive improvement but does not equate improvement. Measurement can help identify opportunities for improvement with actual improvement requiring making changes to health care processes. 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 (Conway, Mostashari, & Clancy, 2013).

Within the literature we found that Makam et al. evaluated trends in the prescription of evidence-based cardiac medications amongst eligible patients. Methods included analyzing over 5,000 patients discharged from various hospitals in central Massachusetts after acute myocardial infarction. Study authors found that between 2001 and 2011, eligible patients receiving a new prescription for aspirin increased from 70% to 88% (2016).

Citations:

1. Conway, P. H., Mostashari, F., & Clancy, C. (2013). The future of quality measurement for improvement and accountability. JAMA, 309(21), 2215–2216. https://doi.org/10.1001/jama.2013.4929

2. Makam, R. C., Erskine, N., McManus, D. D., Lessard, D., Gore, J. M., Yarzebski, J., & Goldberg, R. J. (2016). Decade-Long Trends (2001 to 2011) in the Use of Evidence-Based Medical Therapies at the Time of Hospital Discharge for Patients Surviving Acute Myocardial Infarction. The American journal of cardiology, 118(12), 1792–1797. https://doi.org/10.1016/j.amjcard.2016.08.065

4b2. 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).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

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.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

A recent meta-analysis evaluated the relationship between medication adherence and clinical outcomes amongst ~100,000 patients with stable coronary artery disease. Study authors found that adherence to evidence-based medications, including antiplatelet drugs, was related to a lower risk of all-cause mortality (risk ratio 0.56; 95% CI), lower risk of cardiovascular mortality (risk ratio 0.66; 95% CI), and lower risk of cardiovascular hospitalization/myocardial infarction(risk ratio 0.61; 95% CI) (Du, Cheng, Zhang, Li, & Mei, 2017).

Citation:

Du, L., Cheng, Z., Zhang, Y., Li, Y., & Mei, D. (2017). The impact of medication adherence on clinical outcomes of coronary artery disease: A meta-analysis. European journal of preventive cardiology, 24(9), 962–970. https://doi.org/10.1177/2047487317695628

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)

0465 : Perioperative Anti-platelet Therapy for Patients undergoing Carotid Endarterectomy

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

5a. Harmonization of Related Measures

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 harmonized to the extent possible?

Yes

5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.

The patient population of 0465 is adults undergoing carotid endarterectomy, whereas the patient population of 067 is adults with coronary artery disease.

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.)
Not 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.

No appendix **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.

Bruce Abramowitz, MD, FACC (interventional cardiology; measure implementation)

Karen Alexander, MD (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 (guideline development)

Michael O'Toole, MD (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)

ACC/AHA/ AMA 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 must be equal contributors to the measure development process. In addition, the ACC/AHA/AMA PCPI strive to include on its 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, 2019

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, 2020

Ad.6 Copyright statement: Physician performance measures and related data specifications were developed by the PCPI Foundation, the American College of Cardiology (ACC), and the American Heart Association (AHA) to facilitate quality improvement activities by physicians. These performance measures are not clinical guidelines and do not establish a standard of medical care, and have not been tested for all potential applications. While copyrighted, they can be reproduced and distributed, without modification, for noncommercial purposes, e.g., use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the performance measures for commercial gain, or incorporation of the performance 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 Foundation or the ACC or the AHA. Neither the AMA, ACC, AHA, the PCPI nor its members shall be responsible for any use of these 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, the ACC, the AHA, the PCPI and its members disclaim all liability for use or accuracy of any Current Procedural Terminology (CPT®) or other coding contained in the specifications. CPT® contained in the measures specifications is copyright 2020 American Medical Association.

Ad.7 Disclaimers: Physician performance measures and related data specifications were developed by the PCPI Foundation, the American College of Cardiology (ACC), and the American Heart Association (AHA) to facilitate quality improvement activities by physicians. These performance measures are not clinical guidelines and do not establish a standard of medical care, and have not been tested for all potential applications. While copyrighted, they can be reproduced and distributed, without modification, for noncommercial purposes, e.g., use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the performance measures for commercial gain, or incorporation of the performance 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 Foundation or the ACC or the AHA. Neither the AMA, ACC, AHA, the PCPI nor its members shall be responsible for any use of these 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, the ACC, the AHA, the PCPI and its members disclaim all liability for use or accuracy of any Current Procedural Terminology (CPT®) or other coding contained in the specifications. CPT® contained in the measures specifications is copyright 2020 American Medical Association.

Ad.8 Additional Information/Comments: None.