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 #: 0070

Corresponding Measures: 0070e

De.2. Measure Title: Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)

Co.1.1. Measure Steward: PCPI Foundation

De.3. Brief Description of Measure: Percentage of patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12-month period who also have a prior MI or a current or prior LVEF <40% who were prescribed beta-blocker therapy

1b.1. Developer Rationale: For patients with coronary artery disease (CAD), beta-blockers are recommended for 3 years after myocardial infarction or acute coronary syndrome. Beta-blockers, particularly carvedilol, metoprolol succinate, or bisoprolol which have been shown to reduce risk of death, are recommended indefinitely for patients with CAD and LV systolic dysfunction. These agents have proven efficacy in reducing angina onset and improving the ischemic threshold during exercise. In patients who have suffered an MI, beta-blockers significantly reduce deaths and recurrent MIs. (1) Nonadherence to cardioprotective medications is prevalent among outpatients with CAD and can be associated with a broad range of adverse outcomes, including all-cause and cardiovascular mortality, cardiovascular hospitalizations, and the need for revascularization procedures.(2) This measure is intended to promote beta-blocker usage in select patients with CAD.

References:

1. Fihn SD, Gardin JM, Abrams J, Berra K, Blankenship JC, Dallas AP, Douglas PS, Foody JM, Gerber TC, Hinderliter AL, King SB III, Kligfield PD, Krumholz HM, Kwong RYK, Lim MJ, Linderbaum JA, Mack MJ, Munger MA, Prager RL, Sabik JF, Shaw LJ, Sikkema JD, Smith CR Jr, Smith SC Jr, Spertus JA, Williams SV. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J Am Coll Cardiol 2012;60:e44-164.

2. Gibbons RJ, Abrams J, Chatterjee K, Daley J, Deedwania PC, Douglas JS, Ferguson TB Jr., Fihn SD, Fraker TD Jr., Gardin JM, O'Rourke RA, Pasternak RC, Williams SV. ACC/AHA 2002 guideline update for the management of patients with chronic stable angina: a report of the American College of Cardiology/American Heart

Association Task Force on Practice Guidelines (Committee to Update the 1999 Guidelines for the Management of Patients with Chronic Stable Angina). 2002. Available at: www.acc.org/clinical/guidelines/stable/stable.pdf

S.4. Numerator Statement: Patients who were prescribed beta-blocker therapy

S.6. Denominator Statement: All patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12-month period who also have a prior (within the past 3 years) MI or a current or prior LVEF < 40%

S.8. Denominator Exclusions: Denominator Exceptions:

Documentation of medical reason(s) for not prescribing beta-blocker therapy (e.g., allergy, intolerance, other medical reasons).

Documentation of patient reason(s) for not prescribing beta-blocker therapy (e.g., patient declined, other patient reasons).

Documentation of system reason(s) for not prescribing beta-blocker therapy (e.g., 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 : Group/Practice, Clinician : Individual

IF Endorsement Maintenance – Original Endorsement Date: Aug 10, 2009 **Most Recent Endorsement Date:** Feb 26, 2016

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**

Evidence Summary

- The developer provides a diagram to support the relationship between the process of care (Beta-Blocker Therapy for CAD patients with prior myocardial infarction (MI) or left ventricular systolic dysfunction (LVEF <40%) andreduced risk of death, reduced angina onset, reduced recurrent MIs in patients with a prior MI, and improved ischemic threshold during exercise..
- The developer cites the 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the Diagnosis and Management of Patients With Stable Ischemic Heart Disease, which states: Beta-blocker therapy should be started and continued for 3 years in all patients with normal LV function after MI or ACS. (Class I, Level of Evidence: B); Beta-blocker therapy should be used in all patients with LV systolic dysfunction (EF ≤40%) with heart failure or prior MI, unless contraindicated. (Use should be limited to carvedilol, metoprolol succinate, or bisoprolol, which have been shown to reduce risk of death.) (Class I, Level of Evidence: A).
- For patients with a previous MI, 3 articles support the recommendation (two systematic reviews and one observational study). For LVSD patients, five articles are cited, including 3 randomized controlled trials, one meta-analysis of randomized controlled trials, and one comparative analysis of randomized controlled trials. The overall quality of evidence across studies is not provided. Beta-blockade side effects are discussed and measure harms are noted. An analysis of 75 additional articles since the guidline is provided.

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. Does the Committee agree the evidence basis for the measure has not changed and there is no need for repeat discussion and vote on Evidence?*

Guidance from the Evidence Algorithm

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

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

1b. Gap in Care/Opportunity for Improvement and 1b. Disparities

Maintenance measures – increased emphasis on gap and variation

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

- The developer provided the following registry performance data for 1,100 providers from CMS's PQRS program from January 2016 to December 2016:
  - Number of quality events: 18,558
  - Mean: 0.92
  - Median: 1.00

- Mode: 1.00
- Standard Deviation: 0.14
- Range: 0.93
- Minimum: 0.07
- Maximum: 1.00
- Interquartile Range: 0.13 (1.00–0.88)

- The Registry/QCDR average performance rate reported for the 2018 MIPS benchmark report is 84.2% and standard deviation of 15.2.
- The developer also provided a summary of data from the literature.

Disparities

- No data on disparities from the measure as specified was provided. The developer noted the measure is included in a federal reporting program; however, the program does not provide disparities data – this is required for maintenance of endorsement.
- The developer also provided a summary of disparities data from the literature.

Questions for the Committee:

- Does the performance data continue to warrant a national performance registry measure?
- If no disparities information is provided, 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 and 1b)

1a. Evidence:

- Evidence base largely unchanged. However, there have been meta analyses I believe out of NYU suggesting that beta blocker therapy in CAD may have an "expiration date" after AMI
- Process measure; applies tangentially - prescriptions rather than filled; BB use improves survival; no new evidence
- Evidence sufficient
- The evidence applies directly to an outcome

1b. Performance Gap:

- No disparities data provided. Moderate performance gap of 84%
- Mean was 0.92, Median 1.0, Min 0.07. High mean performance but still room for improvement. Disparities data not provided - cited not available from CMS
- Evidence that a gap exists
- While the mean is greater than .90, the range is significant. A performance gap is present

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.

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

Evaluators: Staff

Questions for the Committee regarding reliability:

- The measure will be considered for endorsement at the clinician group level of analysis and outpatient setting unless additional testing is provided.
- Seek clarification from the developer to determine if the reliability scores are the average reliability for providers with 1+ events and 10+ events.
- Reliability decreased slightly from 0.85 for 1+ events to 0.84 for 10+ events. Does the Committee have any concerns that reliability decreased as the number of events increased?
- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?

Questions for the Committee regarding validity:

- The measure chosen for the correlation analysis (NQF #0066) prescribes ACE and ARB therapy, not beta blocker therapy; however, the developer described two measures that prescribe beta blocker therapy. Does the Committee have any concerns about the discrepancy in the validity testing results?
- Does the Committee have any concerns about the relationship, if any, of patients who also have diabetes (NQF #0066) and the current measure?

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

RATIONALE: Unable to determine validity of the measure due to discrepancy in correlation analysis testing information provided.

Scientific Acceptability: Preliminary Analysis Form

Measure Number: 0070

Measure Title: Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)

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**
2. **Briefly summarize any concerns about the measure specifications.**
    - Levels of analysis and care settings inconsistent with testing provided. The level of analysis (LoA) specified are for individual clinicians and clinician groups. The care settings specified are home care, other, outpatient services, post-acute care, nursing facility visit, and care services in long-term residential facility.
    - The LoA and care settings in the measure specifications must align with testing (clinician group and outpatient services). Additional testing is required for endorsement at the individual clinician level in home care, post-acute care, nursing facility visit, and care services in long-term residential facility.
    - Section 1.5 and 1.6 discuss minimum number of quality reporting events (10) and providers who had 10 or more patients eligible for this measure.
        - The difference between reporting events and patients is not clear.
        - Minimum number of patients and/or reporting events is not included in specifications.

RELIABILITY: TESTING

Submission document: "MIF_xxxx" 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**
   - Reliability testing conducted at clinician group level of analysis only.
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☐ **Yes** ☒ **No**
6. **Assess the method(s) used for reliability testing**
   - Reliability testing conducted at the score level using signal to noise ratio.
   - Providers must have at least 10 eligible reporting events to be included in calculation – this is inconsistent with specifications.

   Submission document: Testing attachment, section 2a2.2
7. **Assess the results of reliability testing**
   - Reliability for 1+ events: 0.85; 10+ events: 0.84. Developer does not state if these results are the average reliability for providers.

   Submission document: Testing attachment, section 2a2.3
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.**
    - Rated moderate due to concerns about measure specifications.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

    Submission document: Testing attachment, section 2b2.

    - Current testing data states providers with minimum (10) number of quality reporting events – this is inconsistent with specifications.
    - Data demonstrates average number of exceptions per provider (1.0); percentage of individuals excluded and frequency distribution of exclusions across providers not included.

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

    Submission document: Testing attachment, section 2b4.

    - Developer repeated performance gap information. NQF guidance states "do not just repeat the information provided related to performance gap in 1b"

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.

    - Missing data analysis not performed – this is required.

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 2b1.2.

- Correlation analysis was conducted for validity testing using the performance measure score on this measure (NQF #0070) and another registry performance measure, NQF #0066: *Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor and Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%)* (PQRS #118) due to similarities in patient population and domain.
- The developer hypothesized a positive association of scores between providers who prescribe **beta blocker therapy** on patients with coronary artery disease seen within a 12 month period and who also have a prior MI or a current or prior LVEF < 40% , and those who prescribe **beta blocker therapy** on patients with a diagnosis of heart failure (HF) with a current or prior left ventricular ejection fraction (LVEF) < 40% within a 12 month period.
- Developer did not provide hypothesized association between this measure and patients also diagnosed with diabetes and prescribed ACE inhibitor or ARB therapy (PQRS #118/NQF #0066).

20. **Assess the results(s) for establishing validity**

Submission document: Testing attachment, section 2b1.3.

- Per the developer, this measure has a moderate positive correlation (0.22) with another evidence-based process of care measure (NQF #0066). However, developer described analyzing the correlation between two measures that prescribe beta blocker therapy on patients with CAD, prior MI, and current and/or prior LVEF <40%.
- NQF #0066 includes patients prescribed ACE inhibitors or ARB therapy not beta blocker therapy.
- Developer did not discuss the relationship, if any, of patients who also have diabetes (NQF #0066) and the current 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)

- See comments above

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.**

- Potential threats to validity that are relevant to the measure not empirically assessed. Unable to determine validity of the measure due to discrepancy in correlation analysis testing information provided.

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

2a. Reliability – Specifications:

- No major concerns
- Concerned with how reliably determination of MI within 3 years can be made; also, accessibility of EF < 40% information in outpatient records (PCP)
- I'm not sure how they document MI within the 3-year time frame
- There are some concerns about measure specs. Reliability is moderate

2a2. Reliability – Testing:

- Reliability score 0.85 (very good); not concerned about 0.01 difference between 1+ events and 10+ events
- Inconsistency between level of analysis and care settings with testing
- No concerns. Reliability was 0.85 for 1+ events and 0.84 for 10+ events

2b1. Validity – Testing:

- Am not sure about testing against ARB in diabetes since this is a different treatment in a different population
- Empirical validity testing results very low Correlation 0.22
- Only weakly correlated with the chosen measures for testing
- Validity was tested by comparing to ACEi/ARB Rx. No concerns

2b4-7. Threat to Validity; Meaningful Differences; Comparability of Performance Scores; Missing Data/No Response:

- Missing data analysis not performed
- The testing dataset can't contain missing data because it would have been rejected from submission. Therefore, we don't know about missing data.
- Meaningful differences in quality do exist. Missing data do not appear to be a threat to the validity of the measure

2b2-3. Other Treats to Validity; Exclusions; Risk Adjustment:

- We need to understand percentage of individuals excluded. I am not sure performance measure can exclude based on patient refusals which this seems to
- Exclusions did not seem to be a problem; There was no risk adjustment
- Exclusions are appropriate. The measure is not risk adjusted.

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.

- Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry).
- All data elements are in defined fields in electronic clinical data (e.g., clinical registry, nursing home MDS, home health OASIS)

Questions for the Committee:

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

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

RATIONALE: Measure requires chart abstraction for registry. Developer did not discuss time and costs associated with abstracting measure; therefore, unable to determine if data captured without undue burden.

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

- Did not discuss time/costs of data extraction
- I would say moderate at best. While this measure has been used since 2016, only 36% of providers in the test data set submitted files that contained all the necessary data elements
- Lack of acceptable information to assess
- I don't identify any problems with feasibility

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**

OR

Planned use in an accountability program? ☐ **Yes** ☐ **No**

Accountability program details

- This measure is currently used in the Merit-based Incentive Payment System (MIPS). The measure was previously used in the Physician Quality Reporting System (PQRS).
- The measure is not currently publicly reported, but data will be available for public reporting in Physician Compare beginning in late 2019.
- The measure is used in the PINNACLE Registry® for internal quality improvement.

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

- The developer states no feedback has been received.

Additional Feedback:

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 developer did not discuss any progress on improvement.

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 did not list any unexpected findings.

Potential harms

- The developer did not list potential harms.

Additional Feedback:

***Questions for the Committee**:*

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

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

RATIONALE: The developer did not discuss any progress on improvement.

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

4a1. Use - Accountability and Transparency:

- Publicly reported - PQRS; developers say no feedback received
- It is publicly reported
- The measure is used in MIPS. PCPI has not received negative feedback

4b1. Usability – Improvement:

- Potential harms not discussed
- Results are used to determine incentive pay for providers for delivering specific interventions; no significant harms perceived
- No information on results over time, no information on potential harms
- No unintended consequences have been identified

Criterion 5: Related and Competing Measures

Related or competing measures

- 0070e : Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)
- 0071 : Persistence of Beta-Blocker Treatment After a Heart Attack
- 0071 : Persistence of Beta-Blocker Treatment After a Heart Attack
- 0083 : Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)
- 0083 : Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)
- 0083e : Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)
- 0117 : Beta Blockade at Discharge
- 0127 : Preoperative Beta Blockade

Harmonization

- Developer states the patient population for this measure is covered by the following NQF-endorsed measures:
  - NQF 0071: Persistence of Beta-Blocker Treatment After a Heart Attack
  - NQF 0083 and 0083e: Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD).

- The denominator specifications for the measures differ where needed based on the differing patient populations. 0071 is intended for use at the health plan level. 0117 is an inpatient/hospital level measure and includes only patients who have undergone isolated CABG surgery. 0127 is also an inpatient/hospital level measure that focuses on administration of beta-blockers prior to isolated CABG surgery. Measure 0070e is the EHR version of this measure and is completely harmonized.

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

- Primary competition is with 0071 BB after heart attack (health plan level)
- 0083 & 0071. Harmonization and consideration of a HF composite measure would be helpful.
- 0070e is the e-measure. 0083 and 0083e and 0117 are related but don't compete.

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: Month/Day/Year

- **Of the XXX NQF members who have submitted a support/non-support choice:**
  - XX support the measure
  - YY do not support the measure

NATIONAL QUALITY FORUM

Measure Information

This document contains the information submitted by measure developers/stewards, but is organized according to NQF's measure evaluation criteria and process. The item numbers refer to those in the submission form but may be in a slightly different order here. In general, the item numbers also reference the related criteria (e.g., item 1b.1 relates to sub criterion 1b).

Brief Measure Information

NQF #: 0070

Corresponding Measures: 0070e

De.2. Measure Title: Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)

Co.1.1. Measure Steward: PCPI Foundation

De.3. Brief Description of Measure: Percentage of patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12-month period who also have a prior MI or a current or prior LVEF <40% who were prescribed beta-blocker therapy

1b.1. Developer Rationale: For patients with coronary artery disease (CAD), beta-blockers are recommended for 3 years after myocardial infarction or acute coronary syndrome. Beta-blockers, particularly carvedilol, metoprolol succinate, or bisoprolol which have been shown to reduce risk of death, are recommended indefinitely for patients with CAD and LV systolic dysfunction. These agents have proven efficacy in reducing angina onset and improving the ischemic threshold during exercise. In patients who have suffered an MI, beta-blockers significantly reduce deaths and recurrent MIs. (1) Nonadherence to cardioprotective medications is prevalent among outpatients with CAD and can be associated with a broad range of adverse outcomes, including all-cause and cardiovascular mortality, cardiovascular hospitalizations, and the need for revascularization procedures.(2) This measure is intended to promote beta-blocker usage in select patients with CAD.

References:

1. Fihn SD, Gardin JM, Abrams J, Berra K, Blankenship JC, Dallas AP, Douglas PS, Foody JM, Gerber TC, Hinderliter AL, King SB III, Kligfield PD, Krumholz HM, Kwong RYK, Lim MJ, Linderbaum JA, Mack MJ, Munger MA, Prager RL, Sabik JF, Shaw LJ, Sikkema JD, Smith CR Jr, Smith SC Jr, Spertus JA, Williams SV. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J Am Coll Cardiol 2012;60:e44-164.

2. Gibbons RJ, Abrams J, Chatterjee K, Daley J, Deedwania PC, Douglas JS, Ferguson TB Jr., Fihn SD, Fraker TD Jr., Gardin JM, O'Rourke RA, Pasternak RC, Williams SV. ACC/AHA 2002 guideline update for the management of patients with chronic stable angina: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1999 Guidelines for the Management of Patients with Chronic Stable Angina). 2002. Available at: www.acc.org/clinical/guidelines/stable/stable.pdf

S.4. Numerator Statement: Patients who were prescribed beta-blocker therapy

S.6. Denominator Statement: All patients aged 18 years and older with a diagnosis of coronary artery disease seen within a 12-month period who also have a prior (within the past 3 years) MI or a current or prior LVEF < 40%

S.8. Denominator Exclusions: Denominator Exceptions:

Documentation of medical reason(s) for not prescribing beta-blocker therapy (e.g., allergy, intolerance, other medical reasons).

Documentation of patient reason(s) for not prescribing beta-blocker therapy (e.g., patient declined, other patient reasons).

Documentation of system reason(s) for not prescribing beta-blocker therapy (e.g., 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 : Group/Practice, Clinician : Individual

IF Endorsement Maintenance – Original Endorsement Date: Aug 10, 2009 **Most Recent Endorsement Date:** Feb 26, 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?

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

nqf_evidence_attachment_0070_FINAL_08APR19.docx

1a.1 <u>For Maintenance of Endorsement:</u> 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

NATIONAL QUALITY FORUM—Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0070
Measure Title: Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)
IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Not Applicable
Date of Submission: 4/9/2019

Instructions

- *Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.*
- *Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.*
- *For composite performance measures:*
  - *A separate evidence form is required for each component measure unless several components were studied together.*
  - *If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.*
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.

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

1a. Evidence to Support the Measure Focus

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

- <u>Outcome</u>: [3] Empirical data demonstrate a relationship between the outcome and at least one healthcare structure,

process, intervention, or service. If not available, wide variation in performance can be used as evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.

- Intermediate clinical outcome: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured intermediate clinical outcome leads to a desired health outcome.
- Process: [5] a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured process leads to a desired health outcome.
- Structure: a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence [4] that the measured structure leads to a desired health outcome.
- Efficiency: [6] evidence not required for the resource use component.
- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- Process measures incorporating Appropriate Use Criteria: See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

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

4. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation (GRADE) guidelines and/or modified GRADE.

5. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

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

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

☐ 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

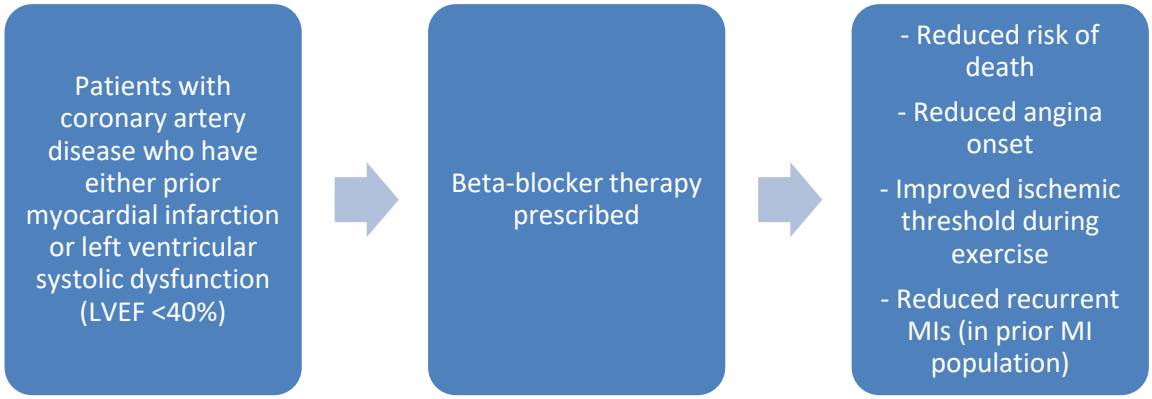
☐ Process: Beta-Blocker Therapy for CAD patients with Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)

☐ 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.



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.) Not applicable.

****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)

☐ 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**	Fihn SD, Gardin JM, Abrams J, et al. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J Am Coll Cardiol 2012;60:e44-164. Available at: http://www.onlinejacc.org/content/60/24/e44?_ga=2.241949633.2111161951.1554060560-1109945046.1554060560
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.	4.4.2.2. BETA-BLOCKER THERAPY: Beta-blocker therapy should be started and continued for 3 years in all patients with normal LV function after MI or ACS. (Class I, Level of Evidence: B) Beta-blocker therapy should be used in all patients with LV systolic dysfunction (EF ≤40%) with heart failure or prior MI, unless contraindicated. (Use should be limited to carvedilol, metoprolol succinate, or bisoprolol, which have been shown to reduce risk of death.) (Class I, Level of Evidence: A)
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	The guideline recommendations refer to 2 distinct patient populations addressed by the measure – 1) patients with a prior (resolved) (within the past 3 years) myocardial infarction and 2) patients with left ventricular systolic dysfunction (LVEF <40%). For the prior MI population, the weight of the evidence in support of the recommendation is rated as Level B. Level B evidence refers to "Data derived from a single randomized trial, or nonrandomized studies." For the LVSD population, the weight of the evidence in support of the recommendation is rated as Level A. Level A evidence refers to "Data derived from multiple randomized clinical trials or meta-analyses."
Provide all other grades and definitions from the evidence grading system	Levels A and B are described above. Level C evidence refers to "Only consensus opinion of experts, case studies, or standard-of-care." Additional details and information about the evidence rating scheme can also be seen table included at the end of this document.
Grade assigned to the **recommendation** with definition of the grade	Both recommendation statements included in section 1a.4.2 have been assigned a Class I recommendation. Class I recommendations refer to " Conditions for which there is evidence and/or general agreement that a given procedure or treatment is beneficial, useful, and effective."

Provide all other grades and definitions from the recommendation grading system	ACCF/AHA guideline methodology categorizes indications as class I, II, or III on the basis of a multifactorial assessment of risk and expected efficacy viewed in the context of current knowledge and the relative strength of this knowledge. These classes summarize the recommendations for procedures or treatments as follows and noted in the table below: Class I: Conditions for which there is evidence and/or general agreement that a given procedure or treatment is useful and effective. Class II: Conditions for which there is conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of a procedure or treatment. • IIa: Weight of evidence/opinion is in favor of usefulness/efficacy • IIb: Usefulness/efficacy is less well established by evidence/opinion. Class III: Conditions for which there is evidence and/or general agreement that the procedure/treatment is not useful/effective e and in some cases may be harmful. • No Benefit- Procedure/Test not helpful or Treatment w/o established proven benefit • Harm- Procedure/Test leads to excess cost w/o benefit or is harmful, and or Treatment is harmful Citation: ACCF/AHA Task Force on Practice Guidelines. Methodology Manual and Policies From the ACCF/AHA Task Force on Practice Guidelines. American College of Cardiology Foundation and American Heart Association, Inc. Cardiosource.com. 2010. Available at: http://assets.cardiosource.com/Methodology_Manual_for_ACC_AHA_Writing_Committees.pdf and http://my.americanheart.org/idc/groups/ahamah-public/@wcm/@sop/documents/downloadable/ucm_319826.pdf
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	Information regarding the total number of studies and type of study designs included in the body of evidence is not available. However, for the prior MI population: the guideline cites 3 articles in support of the recommendation statement. They include 2 systematic reviews including 33 and 82 randomized controlled trials, respectively, dating back to 1980. The third article was an observational study. For the LVSD population: the guideline cites 5 articles in support of the recommendation statement. They include 3 randomized controlled trials, 1 meta-analysis of randomized controlled trials and 1 comparative analysis of randomized controlled trials. Information regarding the overall quality of evidence across studies is not available.

Estimates of benefit and consistency across studies	The guideline does not include an overall estimate of benefit from the body of evidence. However, they do include the following summary information regarding the benefits of beta-blocker therapy, "Decreases in the rate–BP product, AV nodal conduction, and myocardial contractility from beta blockers reduce myocardial oxygen demand, counteracting beta-receptor activity and contributing to a reduction in angina onset, with improvement in the ischemic threshold during exercise and in symptoms. These agents significantly reduce deaths and recurrent MIs in patients who have suffered a MI and are especially effective when a STEMI is complicated by persistent or recurrent ischemia or tachyarrhythmias early after the onset of infarction."
What harms were identified?	The guideline describes the principle adverse effects of beta blockers as fatigue, exercise intolerance, lethargy, insomnia, nightmares, and impotence.

Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	The guidelines reviewed and incorporated relevant new clinical trials published in peer-reviewed journals and articles through December 2011. A Medical Subject Headings (MeSH®) was conducted using the terms "Adrenergic beta-Antagonists"[Mesh] AND "Coronary Artery Disease"[Mesh] to identify articles published after 2011, resulting in 75 articles. The articles that are most relevant to the focus of the body of evidence are described below. 1. Citation: Bangalore S1, Steg G, Deedwania P, Crowley K, Eagle KA, Goto S, Ohman EM, Cannon CP, Smith SC, Zeymer U, Hoffman EB, Messerli FH, Bhatt DL; REACH Registry Investigators. β-Blocker use and clinical outcomes in stable outpatients with and without coronary artery disease. JAMA. 2012 Oct 3;308(13):1340-9. doi: 10.1001/jama.2012.12559. Description: Longitudinal, observational study of patients in the Reduction of Atherothrombosis for Continued Health (REACH) registry who were divided into 3 cohorts: known prior MI (n = 14,043), known CAD without MI (n = 12,012), or those with CAD risk factors only (n = 18,653) to assess the association of β-blocker use with cardiovascular events in stable patients with a prior history of MI, in those with CAD but no history of MI, and in those with only risk factors for CAD. Results: With a median follow-up of 44 months (interquartile range, 35-45 months), event rates were not significantly different in patients with β-blocker use compared with those without β-blocker use for any of the outcomes tested, even in the prior MI cohort (489 [16.93%] vs 532 [18.60%], respectively; hazard ratio [HR], 0.90 [95% CI, 0.79-1.03]; P = .14). In the CAD without MI cohort, the associated event rates were not significantly different in those with β-blocker use for the primary outcome (391 [12.94%]) vs without β-blocker use (405 [13.55%]) (HR, 0.92 [95% CI, 0.79-1.08]; P = .31), with higher rates for the secondary outcome (1101 [30.59%] vs 1002 [27.84%]; odds ratio [OR], 1.14 [95% CI, 1.03-1.27]; P = .01) and for the tertiary outcome of hospitalization (870 [24.17%] vs 773 [21.48%]; OR, 1.17 [95% CI, 1.04-1.30]; P = .01). In the cohort with CAD risk factors only, the event rates were higher for the primary outcome with β-blocker use (467 [14.22%]) vs without β-blocker use (403 [12.11%]) (HR, 1.18 [95% CI, 1.02-1.36]; P = .02), for the secondary outcome (870 [22.01%] vs 797 [20.17%]; OR, 1.12 [95% CI, 1.00-1.24]; P = .04) but not for the tertiary outcomes of MI (89 [2.82%] vs 68 [2.00%]; HR, 1.36 [95% CI, 0.97-1.90]; P = .08) and stroke (210 [6.55%] vs 168 [5.12%]; HR, 1.22 [95% CI, 0.99-1.52]; P = .06). However, in those with recent MI (≤1 year), β-blocker use was associated with a lower incidence of the secondary outcome (OR, 0.77 [95% CI, 0.64-0.92]). Conclusion: Although this observational study found that the use of β-blockers in the populations studied was not associated with a lower risk of composite cardiovascular events, the article received several letters which highlighted 2 primary concerns: 1) the use of an observational study to assess the effectiveness of a drug when large RCTs and meta-analyses already have shown its effectiveness and 2) the study did not distinguish among different types of beta-blockers.

	While there was a focused update of the guideline that supports this measure in 2014, the specific recommendations that support this measure were not included in the update and remain unchanged. We conducted a second search using the Medical Subject Headings (MeSH®) terms "Adrenergic beta-Antagonists"[Mesh] AND "Coronary Artery Disease"[Mesh] to identify articles published after 2015, resulting in 71 articles. None of the articles published in this timeframe were relevant to the body of evidence that supports this measure. **As the measure developer, we would wait until an updated systematic review of the body of evidence is conducted which can confirm or refute the findings of any study published since the guideline was released, considering the full body of evidence available.**

Table 1. Applying Classification of Recommendation and Level of Evidence

ESTIMATE OF CERTAINTY (PRECISION) OF TREATMENT EFFECT	SIZE OF TREATMENT EFFECT: CLASS I *Benefit >>> Risk* Procedure/Treatment **SHOULD** be performed/ administered	CLASS IIa *Benefit >> Risk* *Additional studies with focused objectives needed* **IT IS REASONABLE** to perform procedure/administer treatment	CLASS IIb *Benefit ≥ Risk* *Additional studies with broad objectives needed; additional registry data would be helpful* Procedure/Treatment **MAY BE CONSIDERED**	CLASS III *No Benefit* or CLASS III *Harm* COR III: No benefit — Procedure/Test: Not Helpful — Treatment: No Proven Benefit COR III: Harm — Procedure/Test: Excess Cost w/o Benefit or Harmful — Treatment: Harmful to Patients
LEVEL A Multiple populations evaluated* Data derived from multiple randomized clinical trials or meta-analyses	■ Recommendation that procedure or treatment is useful/effective ■ Sufficient evidence from multiple randomized trials or meta-analyses	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from multiple randomized trials or meta-analyses	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Sufficient evidence from multiple randomized trials or meta-analyses
LEVEL B Limited populations evaluated* Data derived from a single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is useful/effective ■ Evidence from single randomized trial or nonrandomized studies	■ Recommendation in favor of treatment or procedure being useful/effective ■ Some conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation's usefulness/efficacy less well established ■ Greater conflicting evidence from single randomized trial or nonrandomized studies	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Evidence from single randomized trial or nonrandomized studies
LEVEL C Very limited populations evaluated* Only consensus opinion of experts, case studies, or standard of care	■ Recommendation that procedure or treatment is useful/effective ■ Only expert opinion, case studies, or standard of care	■ Recommendation in favor of treatment or procedure being useful/effective ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation's usefulness/efficacy less well established ■ Only diverging expert opinion, case studies, or standard of care	■ Recommendation that procedure or treatment is not useful/effective and may be harmful ■ Only expert opinion, case studies, or standard of care

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.

For patients with coronary artery disease (CAD), beta-blockers are recommended for 3 years after myocardial infarction or acute coronary syndrome. Beta-blockers, particularly carvedilol, metoprolol succinate, or bisoprolol which have been shown to reduce risk of death, are recommended indefinitely for patients with CAD and LV systolic dysfunction. These agents have proven efficacy in reducing angina onset and improving the ischemic threshold during exercise. In patients who have suffered an MI, beta-blockers significantly reduce deaths and recurrent MIs. (1) Nonadherence to cardioprotective medications is prevalent among outpatients with CAD and can be associated with a broad range of adverse outcomes, including all-cause and cardiovascular mortality, cardiovascular hospitalizations, and the need for revascularization procedures.(2) This measure is intended to promote beta-blocker usage in select patients with CAD.

References:

1. Fihn SD, Gardin JM, Abrams J, Berra K, Blankenship JC, Dallas AP, Douglas PS, Foody JM, Gerber TC, Hinderliter AL, King SB III, Kligfield PD, Krumholz HM, Kwong RYK, Lim MJ, Linderbaum JA, Mack MJ, Munger MA, Prager RL, Sabik JF, Shaw LJ, Sikkema JD, Smith CR Jr, Smith SC Jr, Spertus JA, Williams SV. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J Am Coll Cardiol 2012;60:e44-164.

2. Gibbons RJ, Abrams J, Chatterjee K, Daley J, Deedwania PC, Douglas JS, Ferguson TB Jr., Fihn SD, Fraker TD Jr., Gardin JM, O'Rourke RA, Pasternak RC, Williams SV. ACC/AHA 2002 guideline update for the management of patients with chronic stable angina: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Update the 1999 Guidelines for the Management of Patients with Chronic Stable Angina). 2002. Available at: www.acc.org/clinical/guidelines/stable/stable.pdf

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for 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.*

2016 Registry data from the PQRS program was provided to the PCPI by CMS for the purposes of testing the measure. The data are analyzed for the time period January 2016 through December 2016. There were 18,558 quality events included in this reliability testing and analysis. These were the quality events that were associated with providers who had 1 or more quality events eligible for this measure. Based on the sample of 1,100 included providers, the mean performance rate is 0.92, the median performance rate is 1.00 and the

mode is 1.00. The standard deviation is 0.14. The range of the performance rate is 0.93, with a minimum rate of 0.07 and a maximum rate of 1.00. The interquartile range is 0.13 (1.00–0.88). Decile, Performance (1st,0.69; 2th,0.83; 3nd,0.91; 4th,1.0; 5th,1.0; 6th,1.0; 7th,1.0; 8th,1.0; 9th,1.0; 10th,1.0)

Historical PQRS data from the PQRS experience report does not differentiate between EHR and Registry average performance rates. Performance scores over time are for: 2013: 74.2%, 2014: 79.3%, 2015: 85.1%

It should be noted that PQRS was a voluntary reporting program. Overall participation in the program was suboptimal with 72% of eligible professionals using any method to participate in PQRS, in 2016. The performance scores listed above are not consistently derived from a nationally representative sample.

Quality benchmarks for MIPS 2018 were made publicly available in January 2019. As MIPS is a new program, historical PQRS data was used with MIPS eligibility criteria applied in order to create the benchmark. Providers earn points depending what decile of the benchmark they fall into. The Registry/QCDR average performance rate reported in the benchmark report is 84.2% and standard deviation of 15.2. Deciles 3 through 10 are also reported and are as follows: Decile, Performance (3rd, 72.55%-79.06%, 4th, 79.07%-84.43%, 5th, 84.44%-88.51%, 6th, 88.52%-91.17%. 7th, 91.18%-94.28%, 8th, 94.29%-98.03%, 9th, 98.04%-99.99%, 10th, 100.0%. While not made explicit in the publicly available documentation, it is thought that deciles 1 and 2 are not included in the file since providers earn the same amount of points for results in those deciles regardless of performance. No additional data is available at this time.

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.

While rates have improved over time, suboptimal rates of beta-blocker prescriptions among patients with CAD indicated by PQRS data are further evidenced by several recent studies.

A recent observational study using data from the NCDR® PINNACLE registry found that patients with a diagnosis of CAD from January 1, 2013 through May 1, 2014 enrolled in Medicare Advantage (MA) were more likely to be prescribed beta-blockers (80.6% to 78.8%, P<.001) than Medicare FFS patients. MA patients were found to be younger, but also had more co-morbidities than the FFS patients, as well as more likely to receive other guideline recommended therapy such as ACE/ARB and statin therapy. (1)

Maddox and colleagues analyzed data from 2008 through 2010 from the NCDR® PINNACLE Registry®, a national outpatient cardiology practice registry, to assess practice variation of secondary prevention medication prescription among CAD patients. Among eligible patients, beta-blockers were prescribed in 73.3% (63,800/86,999) at their index clinic visit. After inclusion of all visits among eligible patients occurring within the year following the index visit, the rates increased to 77.3%. Among practices, the median prescription rate of beta-blockers for eligible patients at their index clinic visit was 78.4% (range 35.2-100%) and 79.4% (range 46.2-100%) after inclusion of all visits among eligible patients occurring within the year following the index visit.(12)

An earlier study by Chan and colleagues analyzed 2008-9 data from the Pinnacle registry and found slightly higher rates (86.4%) of beta-blocker prescription among CAD patients following an MI. It's important to note that the Chan et al. study examined compliance rates with performance measures among the first 14,000 outpatients enrolled in the PINNACE program as compared to the Maddox et al study which included a larger and more heterogeneous patient and practice population.(23)

References:

1. Figureroa JF, Blumenthal DM, Feyman Y, Frakt AB, Turchin A, Doros G, et al. Differences in management of coronary artery disease in patients with Medicare Advantage vs traditional Fee-for-Service Medicare among cardiology practices. JAMA Cardiology. 2019;4(265-271).

2. Maddox TM, Chan PS, Spertus JA, Tang F, Jones P, Ho PM, Bradley SM, Tsai TT, Bhatt DL, Peterson PN. Variations in coronary artery disease secondary prevention prescriptions among outpatient cardiology practices: insights from the NCDR (National Cardiovascular Data Registry). J Am Coll Cardiol. 2014 Feb 18;63(6):539-46. doi: 10.1016/j.jacc.2013.09.053. Epub 2013 Oct 30.

3. Chan PS, Oetgen WJ, Buchanan D, et al. Cardiac performance measure compliance in outpatients: the American College of Cardiology and National Cardiovascular Data Registry´s PINNACLE (Practice Innovation And Clinical Excellence) program. J. Am. Coll. Cardiol. 2010; 56(1):8–14.

1b.4. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability. *(This is required for 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, the program does not provide disparities data to analyze and report. In Section 1b.5 below, we provide disparities data reported in 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 Chan et al. article cited above conducted a secondary analysis of PINNACLE data for select performance measures to examine whether compliance rates differed by race or sex. The authors found that compliance rates were similar between black and white patients and men and women for all 4 CAD performance measures (including beta-blocker therapy after MI). (1)

A separate analysis was completed using PINNACLE data from 2009 to compare treatment rates by insurance status for 5 quality-of-care indicators for CAD care related to medication treatment. Uninsured patients were less likely to receive ß-blocker therapy after MI as compared with those who had private health insurance (73.3% vs. 80.5%; unadjusted RR=0.91; 95% CI, 0.87-0.95; P<0.001). There were no meaningful differences in treatment rates between patients with public and private insurance. (2)

1. Chan PS, Oetgen WJ, Buchanan D, et al. Cardiac performance measure compliance in outpatients: the American College of Cardiology and National Cardiovascular Data Registry´s PINNACLE (Practice Innovation And Clinical Excellence) program. J. Am. Coll. Cardiol. 2010; 56(1):8–14.

2. Smolderen KG, Spertus JA, Tang F, et al. Treatment Differences by Health Insurance Among Outpatients with Coronary Artery Disease: Insights from the NCDR®. J Am Coll Cardiol. 2013 Mar 12; 61(10): 1069–1075.

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):*

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 form. Additional measure details may be found at: http://www.thepcpi.org/?page=PCPIMeasures

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

This is not an eMeasure **Attachment:**

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

Attachment **Attachment:** NQF0070_I9toI10_conversion-636904075196450947.xlsx

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. For maintenance of endorsement: 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.

No

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

Supporting guidelines and coding value sets included in the measure are reviewed on an annual basis. However, this annual review has not resulted in any changes for this measure.

S.4. Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome) DO NOT include the rationale for the measure.*

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

Patients who were prescribed beta-blocker therapy

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)*

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

Time Period for Data Collection: At least once during the measurement period

Definition:

Prescribed may include prescription given to the patient for beta-blocker therapy at one or more visits in the measurement period OR patient already taking beta-blocker therapy as documented in current medication list.

Beta-blocker therapy:

- For patients with prior LVEF <40%, beta-blocker therapy includes the following: bisoprolol, carvedilol, or sustained release metoprolol succinate.

- For patients with prior MI, beta-blocker therapy includes any agent within the beta-blocker drug class. As of 2015, no recommendations or evidence are cited in current stable ischemic heart disease guidelines for preferential use of specific agents.

Numerator Note: To meet the intent of the measure, the numerator quality action must be performed at the encounter at which the active diagnosis of CAD or history of cardiac surgery proxy is documented.

For Submission Criteria 1, report Quality Data Code, G9189: Beta-blocker therapy prescribed or currently being taken

For Submission Criteria 2, report CPT Category II Code, 4008F: Beta-blocker therapy prescribed or currently being taken

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 who also have a prior (within the past 3 years) MI or a current or prior LVEF < 40%

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.)*

IF an OUTCOME MEASURE, 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

Denominator Note:

The history of cardiac surgery serves as a proxy for a diagnosis of CAD; a diagnosis is not needed if the patient has documented history of cardiac surgery. Only one of the two criteria – a diagnosis of CAD or history of cardiac surgery proxy – is required. To meet the denominator criteria, a patient must have an active diagnosis of CAD (or proxy documented) at the time of the encounter which is used to qualify for the denominator and evaluate the numerator.

The encounter used to evaluate the numerator counts as 1 of the 2 encounters required for denominator inclusion. If the patient meets the CAD diagnosis criterion, the diagnosis needs to be active only at the encounter being evaluated for the numerator action. If the patient meets the proxy of a history of cardiac surgery inclusion criterion, there should be documentation of the proxy at the encounter being evaluated for the numerator action.

Prior Myocardial Infarction (MI) – for Submission Criteria 2, prior MI is limited to those occurring within the past 3 years.

Submission Criteria 1: Patients with left ventricular systolic dysfunction (LVEF <40%)

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, I24.0, I24.1, I24.8, I24.9, I25.10, I25.110, I25.111, I25.118, I25.119, 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

OR

History of cardiac surgery (CPT): 33140, 33510, 33511, 33512, 33513, 33514, 33516, 33517, 33518, 33519, 33521, 33522, 33523, 33533, 33534, 33535, 33536, 92920, 92924, 92928, 92933, 92937, 92941, 92943, 92980, 92981, 92982, 92984, 92995, 92996

AND

Patient encounter during performance period – to be used for numerator evaluation (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

AND

At least one additional patient encounter during 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

WITH OR WITHOUT

Telehealth Modifier: GQ, GT, 95, POS 02

AND

Left ventricular ejection fraction (LVEF) < 40%: G8694

Submission Criteria 2: Patients with a prior (within the past 3 years) myocardial infarction

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, I24.0, I24.1, I24.8, I24.9, I25.10, I25.110, I25.111, I25.118, I25.119, 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

OR

History of cardiac surgery (CPT): 33140, 33510, 33511, 33512, 33513, 33514, 33516, 33517, 33518, 33519, 33521, 33522, 33523, 33533, 33534, 33535, 33536, 92920, 92924, 92928, 92933, 92937, 92941, 92943, 92980, 92981, 92982, 92984, 92995, 92996

AND

Diagnosis for myocardial infarction– includes patient that had a prior (within the past 3 years) myocardial infarction (ICD-10-CM): I21.01, I21.02, I21.09, I21.11, I21.19, I21.21, I21.29, I21.3, I21.4, I21.9, I21.A1, I21.A9, I22.0, I22.1, I22.2, I22.8, I22.9, I24.1, I25.2

AND

Patient encounter during performance period – to be used for numerator evaluation (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

AND

At least one additional patient encounter during 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

WITH OR 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 beta-blocker therapy (e.g., allergy, intolerance, other medical reasons).

Documentation of patient reason(s) for not prescribing beta-blocker therapy (e.g., patient declined, other patient reasons).

Documentation of system reason(s) for not prescribing beta-blocker therapy (e.g., 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 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): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%), exceptions may include medical reason(s) (eg, allergy, intolerance, other medical reasons), patient reason(s) (eg, patient declined, other patient reasons), or system reason(s) (eg, other reasons attributable to the health care system) for not prescribing beta-blocker therapy. 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.

Additional details are as follows:

For Submission Criteria 1 –

Report Quality Data Code, G9190: Documentation of medical reason(s) for not prescribing beta-blocker therapy (eg, allergy, intolerance, other medical reasons).

Report Quality Data Code, G9191: Documentation of patient reason(s) for not prescribing beta-blocker therapy (eg, patient declined, other patient reasons).

Report Quality Data Code, G9192: Documentation of system reason(s) for not prescribing beta-blocker therapy (eg, other reasons attributable to the health care system).

For Submission Criteria 2 –

Append a modifier to CPT Category II Code:

4008F-1P: Documentation of medical reason(s) for not prescribing beta-blocker therapy (eg, allergy, intolerance, other medical reasons).

4008F-2P: Documentation of patient reason(s) for not prescribing beta-blocker therapy (eg, patient declined, other patient reasons).

4008F-3P: Documentation of system reason(s) for not prescribing beta-blocker therapy (eg, 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.)*

This measure is comprised of two submission criteria but is intended to result in one reporting rate. The reporting rate is the aggregate of Submission Criteria 1 and Submission Criteria 2, resulting in a single performance rate. For the purposes of this measure, the single performance rate can be calculated as follows:

Performance Rate = (Numerator 1 + Numerator 2)/ [(Denominator 1 - Denominator Exceptions 1) + (Denominator 2 - Denominator Exceptions 2)]

Calculation algorithm for Submission Criteria 1: Patients with left ventricular systolic dysfunction (LVEF <40%)

1. Find the patients who meet the initial population (i.e., the general group of patients that a set of performance measures is designed to address).

2. From the patients within the initial population criteria, find the patients who qualify for the denominator (i.e., the specific group of patients for inclusion in a specific performance measure based on defined criteria). Note: in some cases the initial population and denominator are identical.

3. From the patients within the denominator, find the patients who meet the numerator criteria (i.e., the group of patients in the denominator for whom a process or outcome of care occurs). Validate that the number of patients in the numerator is less than or equal to the number of patients in the denominator

4. From the patients who did not meet the numerator criteria, determine if the provider has documented that the patient meets any criteria for exception when denominator exceptions have been specified [for this measure: medical reason(s) (e.g., allergy, intolerance, other medical reasons), patient reason(s) (e.g., patient declined, other patient reasons), or system reason(s) (e.g., other reasons attributable to the health care system) for not prescribing beta-blocker therapy]. If the patient meets any exception criteria, they should be removed from the denominator for performance calculation. --Although the exception cases are removed from the denominator population for the performance calculation, the exception rate (i.e., percentage with valid exceptions) should be calculated and reported along with performance rates to track variations in care and highlight possible areas of focus for QI.

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

Calculation algorithm for Submission Criteria 2: Patients with a prior (within the past 3 years) myocardial infarction

1. Find the patients who meet the initial population (ie, the general group of patients that a set of performance measures is designed to address).

2. From the patients within the initial population criteria, find the patients who qualify for the denominator (i.e., the specific group of patients for inclusion in a specific performance measure based on defined criteria). Note: in some cases the initial population and denominator are identical.

3. From the patients within the denominator, find the patients who meet the numerator criteria (i.e., the group of patients in the denominator for whom a process or outcome of care occurs). Validate that the number of patients in the numerator is less than or equal to the number of patients in the denominator

4. From the patients who did not meet the numerator criteria, determine if the provider has documented that the patient meets any criteria for exception when denominator exceptions have been specified [for this measure: medical reason(s) (e.g., allergy, intolerance, other medical reasons), patient reason(s) (e.g., patient declined, other patient reasons), or system reason(s) (eg, other reasons attributable to the health care system) for not prescribing beta-blocker therapy]. If the patient meets any exception criteria, they should be removed from the denominator for performance calculation. --Although the exception cases are removed from the denominator population for the performance calculation, the exception rate (i.e., percentage with valid exceptions) should be calculated and reported along with performance rates to track variations in care and highlight possible areas of focus for QI.

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

S.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. The 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. The 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.

Not applicable.

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 : Group/Practice, Clinician : Individual

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

Home Care, Other, Outpatient Services, Post-Acute Care

If other: Nursing Facility Visit, Care Services in Long-Term Residential Facility

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. The measure is not a composite.

2. Validity – See attached Measure Testing Submission Form

0070r_nqf_testing-attachment_7.1-636849655375188130.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

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 0070

Measure Title: Coronary Artery Disease – Beta Blocker Therapy Prior to MI or LVSD

Date of Submission: 2/8/2019

Type of Measure:

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

Instructions

- Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.

- **For all measures, sections 1, 2a2, 2b1, 2b2, and 2b4 must be completed.**
- **For outcome and resource use measures**, section **2b3** also must be completed.
- If specified for **multiple data sources/sets of specifications** (e.g., claims and EHRs), section **2b5** also must be completed.
- Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b1-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Maximum of 25 pages (*incuding questions/instructions;* minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
- Contact NQF staff regarding questions. Check for resources at Submitting Standards webpage.
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for version 7.1 of the Measure Testing Attachment.

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

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

2b1. Validity testing [11] demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **instrument-based measures (including PRO-PMs) and composite performance measures**, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure; [12]
AND
If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). [13]

2b3. For outcome measures and other measures when indicated (e.g., resource use):
- **an evidence-based risk-adjustment strategy** (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; [14],[15] and has demonstrated adequate discrimination and calibration

OR
- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful** [16] **differences in performance**;
OR
there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses identify the extent and distribution of **missing data** (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

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

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

13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions.

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

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing,(e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

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

Measure Specified to Use Data From: (*must be consistent with data sources entered in S.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:	☐ 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*).

Previously Submitted 2015 GPRO Registry Data

The data source is the Centers for Medicare & Medicaid Service (CMS) PQRS GPRO database.

Current testing data

The data source is Registry data from the Physician Quality Reporting System (PQRS), provided by the Centers for Medicare & Medicaid Services (CMS).

To participate, EPs and Group practices submit performance data such as number of eligible instances (denominator), instances of quality service performed (numerator), number of performance exclusions, reporting rates, and performance rates—in a file format specified by CMS. Data is then summarized at the practice level and includes both EPs participating individually as well as group practices participating through GPRO.

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

Previously Submitted 2015 GPRO Registry Data

The data are for the time period January 2013 – December 31, 2013 and cover the entire United States.

Current testing data

The data are for the time period January 2016 through December 2016 and cover the entire United States. Given the required conversion to ICD-10 in late 2015, the testing was completed on the ICD-10 specified measure.

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*)

Previously Submitted 2015 GPRO Registry Data

The total number of physicians reporting on this measure is 4,124. Of those, 1,724 physicians had all the required data elements and met the minimum number of quality reporting events (10) for inclusion in the reliability analysis. For this measure, 41.8 percent of physicians are included in the analysis, and the average number of quality reporting events is 61.0 for a total of 105,124 events. The range of quality reporting events for 1,724 physicians included is from 1398 to 10. The average number of quality reporting events for the remaining 58.2 percent of physicians who aren't included is 3.6.

Current testing data

We received data from 1,100 providers reporting on this measure through the Registry reporting option for CMS's PQRS in 2016. This dataset reflects a combination of individual provider data and group data and our analysis of the data as a whole is reflected throughout this submission. Of those, 396 providers had all the required data elements and met the minimum number of quality reporting events (10) for a total of 16,555 quality events. For this measure, 36 percent of providers are included in the analysis, and the average number of quality reporting events are 42 for the remaining 16,555. The range of quality reporting events for 396 providers included is from 10 to 431. The average number of quality reporting events for the remaining 64 percent of providers that aren't included is 3.

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*)

Previously Submitted 2015 GPRO Registry Data

There were 105,124 patients included in this testing and analysis. These were the patients that were associated with physicians who had 10 or more patients eligible for this measure.

Current testing data

There were 16,555 patients included in this reliability testing and analysis. These were the patients that were associated with providers who had 10 or more patients eligible for this measure.

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

Previously Submitted 2015 GPRO Registry Data

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

Face Validity

After the measure was fully specified, an expert panel of 12 members was asked to rate their agreement with the following statement:

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

Current testing data

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

Empirical validity testing was conducted with Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%)

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.

Previously Submitted 2015 GPRO Registry Data

This was not captured as part of the testing.

Current testing data

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*)
Data from the both the previous submission and the current data samples were tested using the same reliability testing method.

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.

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.

Reliability is evaluated by averaging over provider specific reliabilities for all providers that meet the minimum number of quality reporting events for the measure. Each provider must have at least 10 eligible reporting events to be included in this calculation.

A reliability equal to zero implies that all the variability in a measure is attributable to measurement error. A reliability equal to one implies that all the variability is attributable to real differences in provider performance. 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*)

Previously Submitted 2015 GPRO Registry Data
For this measure, the reliability at the minimum level of quality reporting events (10) was 0.65. The average number of quality reporting events for physicians included is 61.0. The reliability at the average number of quality reporting events was 0.92.

Current testing data
The reliability above the minimum level of quality reporting events was 0.84. The reliability including providers with less than 10 eligible reporting events is 0.85.

Table 1: Registry Reliability Results

	2015 submission	Current submission
	Reliability	Reliability
1+ events	0.92	0.85
10+ events	0.65	0.84

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?*)

Previously Submitted 2015 GPRO Registry Data

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

Current testing data

This measure has high reliability when evaluated above the minimum level of quality reporting events and high reliability when including providers with less than the minimum level of quality reporting events.

_____

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)*

Previously Submitted 2015 GPRO Registry Data

Face Validity

Face validity of the measure score as an indicator of quality was systematically assessed as follows.
After the measure was fully specified, the expert panel was asked to rate their agreement with the following statement:

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

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

Current testing data – empirical validity correlation testing

For this measure, the PCPI has conducted review and updates to the measure specifications, which satisfy the NQF's ICD-10 Conversion requirements. We are providing the information below to support the three requirements:

- NQF ICD-10-CM Requirement 1: Statement of intent related to ICD-10 CM
  Goal was to convert this measure to a new code set, fully consistent with the original intent of the measure.

- NQF ICD-10-CM Requirement 2: Coding Table
  See attachment in S.2b

- NQF ICD-10-CM Requirement 3: Description of the process used to identify ICD-10 codes
  The PCPI uses the General Equivalence Mappings (GEMs) as a first step in the identification of ICD-10 codes. We then review the ICD-10 codes to confirm their inclusion in the measure is consistent with the measure intent, making additions or deletions as needed. We have an RHIA-credentialed professional on our staff who reviews all ICD-10 coding. For measures included in CMS' Quality Payment Program (QPP), the ICD-10 codes have also been reviewed and vetted by the CMS contractor. Comments received from stakeholders related to ICD-10 coding are first reviewed internally. Depending on the nature of the comment received, we also engage clinical experts to advise us as to whether a change to the specifications is warranted.

Empirical validity correlation testing

Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) (PQRS #118) 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 beta blocker therapy on patients with coronary artery disease seen within a 12 month period and who also have a prior MI or a current or prior LVEF < 40%, and those who prescribe beta blocker therapy on patients with a diagnosis of heart failure (HF) with a current or prior left ventricular ejection fraction (LVEF) < 40% within a 12 month period. Providers included in the analysis met the minimum number of quality reporting events (10) and were cleaned in the same process as the PQRS dataset.

Datasets were reviewed to identify shared providers based on NPI and TIN identifiers. Correlation analysis was then performed to evaluate the association between performance scores of these shared providers.

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

Correlation	Interpretation
> 0.40	Strong
0.20 - 0.40	Moderate
< 0.20	Weak

1. Shortell T. An Introduction to Data Analysis & Presentation. Sociology 712. http://www.shortell.org/book/chap18.html. Accessed July 13, 2018.

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

Previously Submitted 2015 GPRO Registry Data

Face Validity

Our expert panel included 12 members. Panel members were comprised of experts from the AMA-PCPI Measure Advisory Committee. The list of expert panel members is as follows:

Amy Sanders, MD, MS
David Seidenwurm, MD
Dianne V. Jewell, PT, DPT, PhD, CCS, FAACVPR
Janet Sullivan, MD
John Easa, MD, FIPP
Joseph P. Drozda, Jr., MD, FACC
Mark Metersky, MD
Martha J. Radford, MD, FACC, FAHA
Michael O'Dell, MD, MS, MSHA, FAAFP
Richard Bankowitz, MD, MBA, FACP
Scott T. MacDonald, MD
Shannon Sims, MD, PhD

Current testing data – empirical validity correlation testing

Data from the PQRS program were used to perform the correlation analysis for this measure. Data comes from the Registry versions of Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF < 40%) **(PQRS #007)** and Coronary Artery Disease (CAD): Angiotensin-Converting Enzyme (ACE) Inhibitor or Angiotensin Receptor Blocker (ARB) Therapy - Diabetes or Left Ventricular Systolic Dysfunction (LVEF < 40%) **(PQRS #118)**.

Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF < 40%) **(PQRS #007)** demonstrates 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%) **(PQRS #118)**

PQRS #118

Coefficient of correlation = 0.22
P-value = 0.003
Number of shared providers based on NPI and TIN identifiers = 163

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?*)

Previously Submitted 2015 GPRO Registry Data

Face Validity

The results of the expert panel rating of the validity statement were as follows: N = 12; Mean rating = 4.17 and 91.7% of respondents either agree or strongly agree that this measure can accurately distinguish good and poor quality.

Frequency Distribution of Ratings

1 – 1 response (Strongly Disagree)
2 – 0 responses
3 – 0 responses (Neither Agree nor Disagree)
4 – 6 responses
5 – 5 responses (Strongly Agree)

Current testing data
Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF < 40%) has a moderate positive correlation with another evidence-based process of care measure. The correlation is statistically significant at the 90% confidence level and demonstrates the criterion validity of the measure.

_____

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*)

Previously Submitted 2015 GPRO Registry Data
With the information available from the PQRS Registry, we are unable to determine the type of exception reported. However, the exceptions data captured were analyzed to determine frequency and variability across providers.

Current testing data
Exceptions include:
•Documentation of Medical reason(s) for not prescribing beta-blocker therapy.
•Documentation of Patient reason(s) and system reason(s) for not prescribing beta-blocker therapy.
•Documentation of System reason(s) for not prescribing beta-blocker therapy.

Exceptions were analyzed for frequency across providers.

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*)

Previously Submitted 2015 GPRO Registry Data
Amongst the 1,724 physicians with the minimum (10) number of quality reporting events, there were a total of 4,291 exceptions reported. The average number of exceptions per physician in this sample is 2.5. The overall exception rate is 3.9%.

Current testing data
Amongst the 396 providers with the minimum (10) number of quality reporting events, there were a total of 411 exceptions reported. The average number of exceptions per provider in this sample is 1.0. The proportion of exceptions to patients is 0.02.

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*)

Exceptions are necessary to account for those situations when it is not medically appropriate to prescribe beta-blocker therapy. Exceptions are discretionary and the methodology used for measure exception categories are not uniformly relevant across all measures; for this measure, there is a clear rationale to permit an exception for a medical, patient or system reason. Rather than specifying an exhaustive list of explicit reasons for exception for each measure, the measure developer relies on clinicians to link the exception with a specific reason for the decision not to prescribe beta-blocker therapy required by the measure.

Some have indicated concerns with exception reporting including the potential for providers to inappropriately exclude patients to enhance their performance statistics. Research has indicated that levels of exception reporting occur infrequently and are generally valid (Doran et al., 2008), (Kmetik et al., 2011). Furthermore, exception reporting has been found to have substantial benefits: "it is precise, it increases acceptance of [pay for performance] programs by providers, and it ameliorates perverse incentives to refuse care to "difficult" patients." (Doran et al., 2008).

Although this methodology does not require the external reporting of more detailed exception data, the measure developer recommends that providers document the specific reasons for exception in patients' medical records for purposes of optimal patient management and audit-readiness. We also advocate for the systematic review and analysis of each provider's exceptions data to identify practice patterns and opportunities for quality improvement.

Without exceptions, the performance rate would not accurately reflect the true performance of that provider. This would result in an increase in performance failures and false negatives. The additional value of increased data collection of capturing an exception greatly outweighs the reporting burden.

References:
Doran T, Fullwood C, Reeves D, Gravelle H, Roland M. Exclusion of pay for performance targets by English Physicians. New Engl J Med. 2008; 359: 274-84.

Kmetik KS, Otoole MF, Bossley H et al. Exceptions to Outpatient Quality Measures for Coronary Artery Disease in Electronic Health Records. Ann Intern Med. 2011;154:227-234.

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.
Previously Submitted 2015 GPRO Registry Data
Not applicable

Current testing data
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.

Previously Submitted 2015 GPRO Registry Data

Not applicable

Current testing data

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?

Previously Submitted 2015 GPRO Registry Data

Not applicable

Current testing data

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)**

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

Previously Submitted 2015 GPRO Registry Data

Not applicable

Current testing data

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.**

Previously Submitted 2015 GPRO Registry Data

Not applicable

Current testing data

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*)

Previously Submitted 2015 GPRO Registry Data

Not applicable

Current testing data
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*)**:**
Previously Submitted 2015 GPRO Registry Data
Not applicable

Current testing data
Not applicable

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):
Previously Submitted 2015 GPRO Registry Data
Not applicable

Current testing data
Not applicable

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:
Previously Submitted 2015 GPRO Registry Data
Not applicable

Current testing data
Not applicable

2b3.9. Results of Risk Stratification Analysis:
Previously Submitted 2015 GPRO Registry Data
Not applicable

Current testing data
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*)
Previously Submitted 2015 GPRO Registry Data
Not applicable

Current testing data
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*)
Previously Submitted 2015 GPRO Registry Data

Not applicable

Current testing data
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)*

Previously Submitted 2015 GPRO Registry Data
Measures of central tendency, variability, and dispersion were calculated.

Current testing data
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*)

Previously Submitted 2015 GPRO Registry Data
Based on the sample of 1,724 included physicians, the mean performance rate is 0.74, the median performance rate is 0.78 and the mode is 1.00. The standard deviation is 0.20. The range of the performance rate is 1.0, with a minimum rate of 0.00 and a maximum rate of 1.00. The interquartile range is 0.23 (0.65 - 0.88).

Current testing data
Based on the sample of 396 included providers, the mean performance rate is 0.90, the median performance rate is 0.91 and the mode is 1.0. The standard deviation is 0.14. The range of the performance rate is 1.0, with a minimum rate of 0.07 and a maximum rate of 1.0. The interquartile range is 0.18 (0.97–0.79).

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?*)

Previously Submitted 2015 GPRO Registry Data
The range of performance from 0.00 to 1.00 suggests there's clinically meaningful variation across physicians' performance.

Current testing data
The range of performance from 0.07 to 1.0 suggests there's clinically meaningful variation across providers' performance.

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*)

Previously Submitted 2015 GPRO Registry Data

This test was not performed for this measure

Current testing data

This test was not performed for this measure.

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*)

Previously Submitted 2015 GPRO Registry Data

This test was not performed for this measure

Current testing data

This test was not performed for this measure.

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*)

Previously Submitted 2015 GPRO Registry Data

This test was not performed for this measure

Current testing data

This test was not performed for this measure.

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*)

Previously Submitted 2015 GPRO Registry Data

This test was not performed for this measure.

Current testing data

The PQRS 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*)

Previously Submitted 2015 GPRO Registry Data
Data are not available to complete this testing.

Current testing data
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*)

Previously Submitted 2015 GPRO Registry Data
Not Applicable.

Current testing data
The PQRS 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), Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

If other:

3b. Electronic Sources

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

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

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 copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, eg, use by health care providers in connection with their practices. Commercial uses of the Measures require a license agreement between the user and the AMA, (on behalf of the PCPI), ACC or AHA.

Limited proprietary coding is contained in the Measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets.

4. Usability and Use

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

4a. Accountability and Transparency

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

4.1. Current and Planned Use

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

Specific Plan for Use	Current Use (for current use provide URL)

	Public Reporting Physician Compare https://www.medicare.gov/physiciancompare/ Physician Compare https://www.medicare.gov/physiciancompare/ Payment Program Medicare Quality Payment Program Merit-Based Incentive Payment Program (MIPS) https://qpp.cms.gov/ Quality Improvement (Internal to the specific organization) PINNACLE(R) Registry http://cvquality.acc.org/en/NCDR-Home/Registries/Outpatient-Registries.aspx

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. Merit-based Incentive Payment System (MIPS)-Sponsored by the Centers for Medicare and Medicaid Services (CMS)
Prior to 2016, this measure was used for Eligible Providers (EPs) in the Physician Quality Reporting System (PQRS). As of 2017, PQRS has been replaced by the Merit-based Incentive Payment System (MIPS). MIPS is a national performance-based payment program that uses performance scores across several categories to determine payment rates for EPs. MIPS takes a comprehensive approach to payment by basing consideration of quality on a set of evidence-based measures that were primarily developed by clinicians, thus encouraging improvement in clinical practice and supporting advances in technology that allow for easy exchange of information.
2. According to the CY 2019 Quality Payment Program final rule, CMS intends to "make all measures under MIPS quality performance category available for public reporting on Physician Compare in the transition year of the Quality Payment Program, as technically feasible." These measures include those reported via all available submission methods for MIPS-eligible clinicians and groups. Because this measure has been in use for at least one year and meets the minimum sample size requirement for reliability, this measure meets criteria for public reporting. 2018 data will be available for public reporting on Physician Compare in late 2019. This measure is currently included in the downloadable database on the Physician Compare website and is not yet available on individual or group profiles.
3. PINNACLE Registry (URL: http://cvquality.acc.org/en/NCDR-Home/Registries/Outpatient-Registries.aspx)
The PINNACLE Registry® is cardiology´s largest outpatient quality improvement registry, capturing data on coronary artery disease, hypertension, heart failure and atrial fibrillation. The PINNACLE Registry® continues to grow rapidly, with more than 2400 providers representing almost 800 unique office locations across the U.S submitting data to the registry as of the fourth quarter of 2013. As of the fourth quarter of 2013, the registry has more than 13 million patient encounter records. PINNACLE assists practices in understanding and improving care through the production and distribution of quarterly performance reports. These reports, covering all valid patient encounters, detail adherence to 28 cardiovascular clinical measures at the physician, location, and practice levels across coronary artery disease, hypertension, heart failure and atrial fibrillation. All jointly developed ACC/AHA/PCPI performance measures for these topics are reported by the registry.

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?*)

We support the expanded use of this measure in government or other programs, including those intended for accountability or public reporting. The AMA and PCPI do not have any policies that would restrict access to the performance measure specifications or results or that would impede implementation of the measure for any application. We would welcome its implementation in emerging applications such as accountable care organizations (ACO), Medicare Advantage insurance plans or health plans selling on the insurance marketplace.

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.*)

As described above, it is our understanding that CMS is also planning to move towards publicly reporting physician data via Physician Compare. This measure is currently included in the downloadable database on the Physician Compare website and is not yet available on individual or group profiles. Also, although the measure is currently in use, we support expanded use of this measure in government or other programs, including those intended for accountability or public reporting.

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.

The PCPI measure development and maintenance process is a rigorous, evidence-based process that has been refined and standardized since the PCPI's inception in 2000. Throughout its tenure, the PCPI has conducted its measure development and maintenance process with strict adherence to several key principles, including the following which underscore the role those being measured have played in the development and maintenance process and in providing feedback based on measure implementation:

Collaborative Approach to Measure Development

PCPI measures are developed and maintained through cross-specialty, multi-disciplinary technical expert panels. Representatives of relevant clinical specialties are invited to participate in our expert panels to advise us throughout the measure development process and as questions arise during measure implementation. Additionally, other health care providers and stakeholders participate in our panels as equal contributors to the measure development process. The PCPI also strives to include on its panels individuals representing the perspectives of patients, consumers, private health plans, and employers. Liaisons from key measure development organizations, including The Joint Commission and NCQA, at times participate in the PCPI's measure development process to ensure measure harmonization. Measure methodologists and coding and informatics experts are also considered important members of the expert panel. This broad-based approach to measure development maximizes the input from those being measured and other stakeholders to develop evidence-based, feasible and clinically meaningful measures.

Public Comment Period

Input from a wide range of stakeholders is integral to the measure development process. To invite other perspectives and expertise beyond the expert panels and particularly from those providers and facilities that will implement these measures, the PCPI submits the measures for public comment. All measures are released for a 30-day public and PCPI member comment period. All comments are reviewed by the technical expert panel to determine whether measure modifications are needed based on comments received.

Feedback Mechanisms

The PCPI has a dedicated mechanism set up to receive measure-related comments and questions from implementers. As comments and questions are received, they are shared with appropriate staff for follow up. If comments or questions require expert input, these are shared with the PCPI's technical expert panels to determine if measure modifications may be warranted. Additionally, for PCPI measures included in federal

reporting programs, there is a system that has been set up to elicit timely feedback and responses from PCPI staff in consultation with technical expert panel members, as appropriate.

Feasibility Assessments

The PCPI solicits feedback on measure feasibility in the following domains: data availability, data accuracy, data standards, and workflow to guide future modifications to the measure. During this process, we may receive recommendations to improve the experience of those implementing and reporting on this measure and we follow up on any questions or concerns received by those completing the feasibility assessment. Doing so addresses any issues with interpretation and serves as an important step in the measure development process.

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.

See description in Section 4a2.1.1 above.

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.

As described in Section 4a2.1.1, the PCPI invites feedback through various mechanisms. We obtain input from our topic-specific technical expert panels during the measure development and during the annual maintenance process. Additionally, the PCPI obtains feedback via an online public comment and an email-based process set up to receive measure inquiries from implementers.

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

We have received no feedback from those being measured 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 have resulted in 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 provider 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 of patients diagnosed with coronary artery disease. CMS data report an improvement in performance rates in the last 6 years. However, performance rates represent but one facet of the quality improvement process.

While the PCPI creates measures with an ultimate goal of improving the quality of care, measurement is a mechanism to drive improvement but does not equate with improvement. Measurement can help identify opportunities for improvement with actual improvement requiring making changes to health care processes and/or structure. In order to promote improvement, quality measurement systems need to provide feedback to front-line clinical staff in as close to real time as possible and at the point of care whenever possible. (1)

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

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 have not received reports of unexpected findings resulting from the implementation of this measure. The PCPI has various mechanisms in place for measure users to provide feedback and to identify issues related to the maintenance and implementation of this measure. We convene several topic-specific technical expert panels comprised of various stakeholders including those being measured to advise us regarding any unexpected findings and actions that can be taken to mitigate them.

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

As the prescription of beta-blockers for patients with CAD who have had a prior myocardial infarction or who have LVEF <40% is part of the pharmacotherapy piece of guideline directed medical therapy (along with prescription of antiplatelet therapy and prescription of ACE inhibitor or ARB or ARNI therapy for those for whom it is recommended), it could be anticipated that rates of prescribing these therapies would show improvement as well.

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)

0070e : Coronary Artery Disease (CAD): Beta-Blocker Therapy-Prior Myocardial Infarction (MI) or Left Ventricular Systolic Dysfunction (LVEF <40%)

0071 : Persistence of Beta-Blocker Treatment After a Heart Attack

0071 : Persistence of Beta-Blocker Treatment After a Heart Attack

0083 : Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)

0083 : Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)

0083e : Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD)

0117 : Beta Blockade at Discharge

0127 : Preoperative Beta Blockade

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.

Measure 0070 addresses a patient population of patients with CAD and either a recent prior MI or LVSD. This patient population is also covered in part by the following NQF-endorsed measures: NQF 0071: Persistence of Beta-Blocker Treatment After a Heart Attack and NQF 0083 and 0083e: Heart Failure (HF): Beta-Blocker Therapy for Left Ventricular Systolic Dysfunction (LVSD). The specifications are harmonized to the extent possible. As a result, the denominator specifications for the measures differ where needed based on the differing patient populations. Additionally, NQF 0071 is intended for use at the health plan level. NQF 0117 is an inpatient/hospital level measure and includes only patients who have undergone isolated CABG surgery. NQF 0127 is also an inpatient/hospital level measure that focuses on administration of beta-blockers prior to isolated CABG surgery. Measure 0070e is the EHR version of this measure and is completely harmonized.

5b. Competing Measures

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

OR

Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):

Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)

Appendix

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

Attachment:

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): PCPI Foundation

Co.2 Point of Contact: Samantha, Tierney, samantha.tierney@thepcpi.org, 312-224-6071-

Co.3 Measure Developer if different from Measure Steward: PCPI Foundation

Co.4 Point of Contact: Kerri, Fei, kerri.fei@thepcpi.org, 312-224-6070-

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.

PCPI measures are developed and maintained under the aegis of topic-specific technical expert panels (TEPs). The PCPI TEPs are comprised of clinicians and other healthcare professionals representing medical specialty societies and other stakeholders. The TEPs provide clinical expertise as well as advise on methodologic questions and review the measures annually to ensure accuracy and adherence to the most current evidence.

Cardiovascular Technical Expert Panel

Sarah J. Goodlin MD, FACC, FAAHPM (Co-Chair)

Ileana L. Piña MD, MPH (Co-Chair)

Donald E. Casey MD, MPH, MBA

Ted Ganiats MD

Kathleen L. Grady PhD, RN, FAAN

Richard Hellman MD, FACP, FACE

Tony Hermann

Denise M. Kolanczyk PharmD, BCPS-AQ Cardiology

Frederick A. Masoudi MD, MSPH

Joseph V. Messer MD, MACC

David S. Nilasena MD, MSPH, MS

Stephen D. Persell MD, MPH

Paul D. Rockswold MD, MPH, FAAFP

Nancy K. Sweitzer MD, PhD

Carmen M. Terzic MD, PhD

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: 2019

Ad.4 What is your frequency for review/update of this measure? Supporting guidelines and specifications for this measure are reviewed on an annual basis.

Ad.5 When is the next scheduled review/update for this measure? 2020

Ad.7 Disclaimers: The Measure is not a clinical guideline, does not establish a standard of medical care, and has not been tested for all potential applications.

The Measure, while copyrighted, 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 Measure for commercial gain, or incorporation of the Measure into a product or service that is sold, licensed or distributed for commercial gain.

Commercial uses of the Measure require a license agreement between the user and the American Medical Association (AMA), American College of Cardiology (ACC) or the American Heart Association (AHA). Neither the AMA, nor ACC, nor AHA, nor the former AMA-convened Physician Consortium for Performance Improvement(R) (AMA-PCPI), nor the PCPI (R) Foundation (PCPI[R]), nor their members shall be responsible for any use of the Measure.

AMA and PCPI encourage use of the Measure by other health care professionals, where appropriate.

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

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

Ad.8 Additional Information/Comments: