THE NATIONAL QUALITY FORUM

COMPOSITE MEASURE SUBMISSION FORM
Version 4.1 January 2010

This form will be used by stewards to submit **composite** measures and by reviewers to evaluate the measures.

Measure Stewards: Check with NQF staff before using this form. Complete all non-shaded areas of the form. All requested information should be entered directly into this form. The information requested is directly related to NQF's composite measure evaluation criteria and will be used by reviewers to determine if the evaluation criteria have been met. The specific relevant subcriteria language is provided in a Word comment within the form and will appear if your cursor is over the highlighted area (or in balloons).

The measure steward has the opportunity to identify and present the information that demonstrates the measure meets the criteria. Additional materials will only be considered supplemental. Do not rely solely on materials provided at URLs or in attached documents to provide measure specifications or to demonstrate meeting the criteria. If supplemental materials are provided, be sure to indicate specific page numbers/ web page locations for the relevant information (web page links preferred).

For questions about completing this form, contact the project director at 202-783-1300. Please email this form to the appropriate contact listed in the corresponding call for measures.

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

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

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

Evaluation ratings of the extent to which the criteria are met
C = Completely (unquestionably demonstrated to meet the criterion)
P = Partially (demonstrated to partially meet the criterion)
M = Minimally (addressed BUT demonstrated to only minimally meet the criterion)
N = Not at all (NOT addressed; OR incorrectly addressed; OR demonstrated to NOT meet the criterion)
NA = Not applicable (only an option for a few subcriteria as indicated)

(for NQF staff use) **NQF Review #:** **NQF Project:**
De.1 **Title of Measure:** Composite Measure of Hospital Quality for Acute Myocardial Infarction (AMI)
De.2 **Brief description of measure** (*including type of score, measure focus, target population, time, e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year*): A composite measure of in-hospital process- and outcome-of-care for Acute Myocardial Infarction (AMI) patients.
De.3 **Type of Measure:** ☐ Composite with component measures combined at patient-level (e.g., all-or-none) ☒ Composite with component measures combined at aggregate-level
Select the most relevant priority area(s), quality domain(s), and consumer need(s). De.4 **National Priority Partners Priority Area** ☐ patient and family engagement ☐ population health ☐ safety ☒ care coordination ☐ palliative and end of life care ☐ overuse De.5 **IOM Quality Domain** ☒ effectiveness ☐ efficiency ☐ equity ☐ patient-centered ☐ safety ☒ timeliness

De.6 **Consumer Care Need** ☒ Getting Better ☐ Living With Illness ☐ Staying Healthy

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	**NQF Staff**
A. The measure is in the public domain or an intellectual property agreement (measure steward agreement) is signed. *Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.* A.1 **Do you attest that the measure steward holds intellectual property rights to the measure and the right to use any aspects of the measure owned by another entity (e.g., component measures, risk model, code set)?** ☒ Yes A.2 **Measure Steward Agreement** ☐ Signed and Submitted **OR** ☒ Government entity-public domain (If measure steward agreement not signed for non-government entities, do not submit) A.3 **Please check if either of the following apply:** ☐ Proprietary Measure ☐ Proprietary Complex Measure w/fees	**A** Y☐ N☐
B. The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. B.1 ☒ Yes (If no, do not submit)	**B** Y☐ N☐
C. The intended use of the measure includes both public reporting and quality improvement. C.1 **Purpose:** ☒ Public reporting ☒ Internal quality improvement C.2 ☒ Accountability ☐ Accreditation ☐ Payment incentive ☐ Other, describe: (If not intended for both public reporting and quality improvement, do not submit)	**C** Y☐ N☐
D. The requested measure submission information is complete. Composite measures should be fully developed and tested so that all the evaluation criteria have been addressed and information needed to evaluate the measure is provided. D.1 **Testing:** ☒ Fully developed and tested (If composite measure not tested, do not submit) D.2 **Have NQF-endorsed measures been reviewed to identify if there are similar or related measures?** ☒ Yes (If no, do not submit) *If there are similar or related measures, be sure to address items 3b and 3c with specific information.* ►**Is all requested information entered into this form?** ☒ Yes (If no, do not submit)	**D** Y☐ N☐
De.7 If component measures of the composite **are aggregate-level measures, all** must be either NQF-endorsed or submitted for consideration for NQF endorsement (*check one*) ☒ All component measures are NQF-endorsed measures ☐ Some or all component measures are not NQF-endorsed and have been submitted using the online measure submission tool (If not, do not submit)	Y☐ N☐
(for NQF staff use) **Have all conditions for consideration been met?** Staff Notes to Steward (if submission returned):	**Met** Y☐ N☐
Staff Notes to Reviewers (issues or questions regarding any criteria): **Staff Reviewer Name(s):**	

Comment [KP1]: The individual measures included in the composite or subcomposite measures must be either:
NQF-endorsed;
OR
assessed to have met the individual measure evaluation criteria as the first step in evaluating the composite measure.
(This does not apply to subscales of a scale/instrument that cannot be used independently of the total scale.)

TAP/Workgroup Reviewer Name:	
Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. *Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria.* (composite measure evaluation criteria)	**Eval**
(for NQF staff use) **Specific NPP goal**:	
1d. Purpose/objective of the Composite **1d.1 Describe the purpose/objective of the composite measure:** This measure was designed specifically for use in the Centers for Medicare & Medicaid Services' (CMS) public reporting efforts for measures used in CMS' Hospital Inpatient Quality Reporting Program (formerly RHQDAPU). This program is required to publicly report the adopted measures in particular focus areas related to the quality of hospital inpatient care. The number of measures in the program has expanded considerably, and in the latest inpatient prospective payment system (IPPS) rule, CMS further expanded the measure set to include 60 measures over the next few years. The volume of measures presents a challenge for the public reporting requirement of the program to present this information in a manner that is understandable and useful. The primary objective of this measure is to summarize the measures for the Acute Myocardial Infarction (AMI) focus area into a single composite that is useful, understandable, and acceptable to a wide range of stakeholders. As a result, it is a so-called formative measure. Further discussoin of the construction of formative composite measures appears in Appendix B. Specifically, this measure summarizes both clinical process- and outcome-of-care indicators associated with the treatment of AMI and reported for CMS' Hospital Inpatient Quality Reporting Program. Measures were adopted for this program because, based on a consensus process, they were deemed to be indicators of well-coordinated, high-quality care in the hospital inpatient setting for the clinical condition of interest. In addition, CMS sought an approach to composite methodology that was flexible and adaptable to changes in the sets of measures and clinical conditions included now and in the future of the Hospital Inpatient Quality Reporting program. A condition-specific composite is useful for three reasons. First, in any composite, information from a number of component measures is summarized into a single measure for more effective communication. Second, in a condition-specific composite, the component measures are aggregated at a level that is relevant to both consumers and providers. A condition-specific composite strikes a useful balance between creating one global hospital measure, which might not be relevant to individual consumers or providers with specific needs or practice spheres, and offering only the component measures, which some stakeholders could find overwhelming or contradictory and thus unhelpful. Third, condition-specific composite measures respond simply and directly to a key patient-centered question: "Which hospital should I go to, given my condition?" Moreover, the use of condition-specific composite measures permits disease-specific care teams and their management within hospitals to answer the following question: "Overall, how well is our system serving patients with this condition?" As background, the Hospital Inpatient Quality Reporting Program was initially developed as a result of the Medicare Prescription Drug, Improvement and Modernization Act (MMA) of 2003. Section 5001(a) of Pub. 109-171 of the Deficit Reduction Act (DRA) of 2005 set out new requirements for the program, which built on the ongoing voluntary Hospital Quality Initiative. The Hospital Inpatient Quality Reporting Program is the main effort of CMS to communicate hospital-level quality to patients and providers. **1d.2 Describe the quality construct used in developing the composite:** The composite measure of quality of hospital care for AMI aims to be a comprehensive indicator of hospital performance that will be of special value to consumers as a summary means of evaluating alternative hospitals. The quality construct is thus formative rather than reflective in nature. At present, CMS publishes seven individual process-of-care indicators and two outcome-of-care indicators meant to capture the quality of hospital care provided to patients with AMI. The proposed composite combines these in the form of process- and outcome-of-care domains. CMS developed the composite measure to achieve the following goals for reporting hospital quality measures composite methodology:	**1d** **C**☐ **P**☐ **M**☐ **N**☐

Comment [KP2]: **1d.** The purpose/objective of the composite measure and the construct for quality are clearly described.

- Summarize measures on Hospital Compare in a single, useful, condition-specific composite
- Produce composite values that show differences in hospital performance that are clinically and statistically meaningful and reflect true underlying differences in quality
- Enable the calculation of results for most hospitals
- Employ a method that accommodates changes in the set of measures on Hospital Compare and can be used for multiple conditions
- Employ a method that is relatively simple, so hospitals can duplicate results

These goals can be achieved by a method that is consistent with that of other widely used composites; in this case the method used for the Agency for Healthcare Research and Quality (AHRQ) composites. The National Quality Forum (NQF) has endorsed those composites and CMS, states, and other organizations use them widely.

The current Hospital Inpatient Quality Reporting Program construct domains focus on diseases important to the Medicare population: AMI, Heart Failure (HF), and Pneumonia (PN), and on quality indicators related to the Surgical Care Improvement Project (SCIP). The first three have separate sub-composites in processes- and outcomes-of-care. This system of domains and sub-composites allows addition or removal of measures without changes in methodology or weighting, as well as the publication or analysis of separate process and outcome composites within a condition if desired.

In the development of this composite, certain methodological decisions were made to satisfy the policy goals outlined above. First, we entered individual measures as values, rather than ranks, to reduce the likelihood that very small differences in absolute performance lead to large differences in ranking composite scores. Second, we imputed values for missing indicators so that the composite would define as many hospitals as possible. Third, we adjusted individual measures for reliability, a process that leads to a more accurate measure of true underlying performance and avoids extreme values for small hospitals due to random variation. Lastly, we used denominator weighting so that the composite places more weight on measures that are reported for relatively more patients nationally. In Table 1d.2.1 of Appendix A, we present the mapping between CMS' policy goals and methodological decisions in tabular form.

1e. Components and conceptual construct for quality

1e.1 Describe how the component measures/items are consistent with and representative of the quality construct:

As indicated previously, this composite measure is primarily a formative summary of the measures on Hospital Compare. Thus, the composite includes all measures associated with this condition that are reported on Hospital Compare.

That said, measures were adopted for the Hospital Inpatient Quality Reporting Program because, based on a consensus process, they were deemed to be indicators of well-coordinated, high-quality care in the hospital inpatient setting for the clinical condition of interest. For the AMI, HF, and PN composite measures, the measures that make up the composite include both process- and outcome-of-care indicators; the SCIP composite is made up of process-of-care indicators only.

The composite includes both process- and outcome-of-care indicators, because both types of indicators contain information about quality of care. While it is not possible to directly assess an abstract concept such as quality of care, process-of-care indicators that evaluate whether certain best practices were executed provide critical insight into a hospital's care delivery system. For example, for the AMI composite measure, the component process-of-care indicators evaluate whether a patient received:

- Aspirin on arrival
- Aspirin at discharge
- ACE Inhibitor or ARB for Left Ventricular Systolic Dysfunction (LVSD)
- Smoking Cessation advice/counseling
- Beta Blocker at discharge
- Fibrinolytic medication within 30 minutes of arrival
- PCI within 90 minutes of arrival

These NQF-endorsed process-of-care indicators represent established best practices for AMI care (1, 2) and

1e
C☐
P☐
M☐
N☐

Comment [KP3]: 1e. The component items/measures (e.g., types, focus) that are included in the composite are consistent with and representative of the conceptual construct for quality represented by the composite measure. Whether the composite measure development begins with a conceptual construct or a set of measures, the measures included must be conceptually coherent and consistent with the purpose.

CMS adopted them for the Hospital Inpatient Quality Reporting Program initiative. As standards in clinical practice evolve, additions or changes to these component measures are likely to follow, as well as developing expansions into other conditions and disease states.

In addition to reflecting current clinical guidelines, studies have shown a clear relationship between execution of these practices and decreased mortality for AMI patients (3-5), one of the two outcome-of-care indicators also included in the proposed AMI composite measure. The two AMI outcome-of-care component measures are: 1) 30-day risk-standardized mortality and 2) 30-day risk standardized all-cause readmission. Similar to the process-of-care indicators, these two outcome-of-care indicators are NQF-endorsed and part of CMS' Hospital Inpatient Quality Reporting Program initiative. They directly report the rate of the undesired outcomes (mortality or readmission) that AMI patients at a given hospital experience, and therefore might be critical to understanding the quality of care received.(i)

The combination of these component indicators, each of which is intended to indicate the quality of care received for a subset of patients (that is, AMI, HF, PN, or SCIP), ultimately serves to deliver a single, useful, condition-specific summary for consumer use.

Citations

1. Anderson, JL, Adams, CD, Antman, EM, et al. ACC/AHA 2007 guidelines for the management of patients with unstable angina/non-ST-elevation myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Revise the 2002 Guidelines for the Management of Patients with Unstable Angina/Non-ST-Elevation Myocardial Infarction): developed in collaboration with the American College of Emergency Physicians, American College of Physicians, Society for Academic Emergency Medicine, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons. J Am Coll Cardiol. 2007; 50:e1-157.
2. Antman, EM, Anbe, DT, Armstrong ,PW, et al. ACC/AHA guidelines for the management of patients with ST-elevation myocardial infarction: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1999 Guidelines for the Management of Patients with Acute Myocardial Infarction). 2004.
3. Smith, SC, Allen, J, Blair, SN, et al. AHA/ACC guidelines for secondary prevention for patients with coronary and other atherosclerotic vascular disease: 2006 update. J Am Coll Cardiol. 2006; 47:2130-9.
4. Flather, MD, Yusuf, S, Kober, L, et al. Long-term ACE-inhibitor therapy in patients with heart failure or left-ventricular dysfunction: A systematic overview of data from individual patients. ACE-Inhibitor Myocardial Infarction Collaborative Group. Lancet. 2000; 355(9215):1575-1581.
5. Yusuf, S, Wittes, J, & Friedman, L. Overview of results of randomized clinical trials in heart disease. I. Treatments following myocardial infarction. JAMA. 1988; 260(14):2088:2093

Footnotes

i. In order to align these two indicators with the process-of-care indicators, which report desired, rather than undesired, outcomes, each outcome-of-care indicator is subtracted from 100. This produces two desired outcomes - lack of 30-day mortality and lack of 30-day readmission - which are incorporated into the composite measure.

If the component measures are combined at the patient level, complete 1a, 1b, and 1c.

If the component measures are combined at the aggregate level, skip to criterion 2, *Scientific Acceptability of Measure Properties* (individual measures are either NQF-endorsed or submitted individually).

1a. High Impact

1a.1 Demonstrated high impact aspect of healthcare (*Select the most relevant*)

☐ affects large numbers ☐ frequently performed procedure ☐ leading cause of morbidity/mortality
☐ high resource use ☐ severity of illness ☐ patient/societal consequences of poor quality
☐ other, describe: **1a.2**

1a.3 Summary of Evidence of High Impact:

1a.4 Citations for Evidence of High Impact:

1a
H☐
M☐
L☐
N☐

1b. Opportunity for Improvement

1b.1 Briefly explain benefits (improvements in quality) envisioned by use of this measure:

1b
H☐

Comment [KP4]: 1a. The measure focus addresses:
- a specific national health goal/priority identified by NQF's National Priorities Partners; OR
- a demonstrated high impact aspect of healthcare (e.g., affects large numbers, leading cause of morbidity/mortality, high resource use (current and/or future), severity of illness, and patient/societal consequences of poor quality).

Comment [KP5]: 1b. Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating considerable variation, or overall poor performance, in the quality of care across providers and/or population groups (disparities in care).

1b.2 Summary of data demonstrating performance gap (*variation or overall poor performance across providers*)**:** **1b.3 Citations for data on performance gap:** **1b.4 Summary of Data on disparities by population group:** **1b.5 Citations for data on Disparities:**	**M☐ L☐ N☐**
1c. Evidence-based **1c.1 Relationship to Outcomes** (*For non-outcome measures, briefly describe the relationship to desired outcome. For outcomes, describe why it is relevant to the target population.)* **1c.2 Type of Evidence** *(Check all that apply)* ☐ Cohort study ☐ Evidence-based guideline ☐ Expert opinion ☐ Meta-analysis ☐ Observational study ☐ Randomized controlled trial ☐ Systematic synthesis of research ☐ Other (*Please describe*): **1c.3** **1c.4 Summary of Evidence** *as described above for type of measure; for outcomes, summarize any evidence that healthcare services/care processes influence the outcome*)**:** **1c.5 Rating of strength/quality of evidence** (*also provide narrative description of the rating and by whom*) **1c.6 Method for rating evidence:** **1c.7 Summary of Controversy/Contradictory Evidence:** **1c.8 Citations for Evidence** (*other than guidelines*) **1c.9 Quote the Specific guideline recommendation** (*including guideline number and/or page number*) **1c.10 Clinical Practice Guideline Citation:** **1c.11 National Guideline Clearinghouse or other URL:** **1c.12 Rating of strength of recommendation** (*also provide narrative description of the rating and by whom*) **1c.13 Method for rating strength of recommendation** (*If different from USPSTF system, also describe rating and how it relates to USPSTF)***:** **1c.14 Rationale for using this guideline over others:**	**1c** **H☐ M☐ L☐ N☐**
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	**1**
Steering Committee: Was the threshold criterion, *Importance to Measure and Report***, met?** **Rationale:**	**1** **Y☐ N☐**
2. 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. (composite measure evaluation criteria)	**Eval**
2a. COMPOSITE MEASURE SPECIFICATIONS	
In the future, NQF will require measure stewards to provide a URL link to a web page where current detailed specifications can be obtained? **S.1** Do you have a web page where current detailed measure specifications can be obtained? Upon endorsement, the proposed measure specifications will be posted on the Hospital Compare website: http://www.hospitalcompare.hhs.gov/	**2a-specs** **C☐ P☐ M☐**

Comment [KP6]: 1c. The measure focus is:
- an outcome (e.g., morbidity, mortality, function, health-related quality of life) that is relevant to, or associated with, a national health goal/priority, the condition, population, and/or care being addressed;

OR
- if an intermediate outcome, process, structure, etc., there is **evidence** that supports the specific measure focus as follows:
  - Intermediate outcome - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
  - Process - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
  - Structure - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
  - Patient experience - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
  - Access - evidence that an association exists between access to a health service and the outcomes of, or experience with, care.

Efficiency - demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

S.2 If yes, provide web page URL: http://www.hospitalcompare.hhs.gov/

N☐

2a. Precisely Specified

2a.0.1 Components of the Composite (*List the components, i.e., domains/sub-composites, individual measures. If component measures are NQF-endorsed, include NQF measure number*; if not NQF-endorsed, provide date of submission to NQF)

HOSPITAL PROCESS-OF-CARE INDICATORS

1. Percent of AMI Patients Given Aspirin on Arrival (NQF #0132; Endorsed May 9, 2007)
2. Percent of AMI Patients Given Aspirin at Discharge (NQF #0142; Endorsed May 9, 2007)
3. Percent of AMI Patients Given ACE Inhibitor or ARB for LVSD (NQF #0137; Endorsed May 9, 2007)
4. Percent of AMI Patients Given Smoking Cessation Advice/Counseling (NQF #0027; Endorsed May 1, 2006)
5. Percent of AMI Patients Given Beta Blocker at Discharge (NQF #0160; Endorsed May 9, 2007)
6. Percent of AMI Patients Given Fibrinolytic Medication within 30 Min. of Arrival (NQF #0164; Endorsed May 9, 2007)
7. Percent of AMI Patients Given PCI within 90 Min. of Arrival (NQF #0163; Endorsed May 9, 2007)

HOSPITAL OUTCOME-OF-CARE INDICATORS

1. AMI 30-day Risk-Standardized Mortality (NQF #0230; Endorsed May 9, 2007)
2. AMI 30-day Risk-Standardized Readmission (NQF #0505; Endorsed Oct. 28, 2008)

If the composite measure cannot be specified with a numerator and denominator, please consult with NQF staff.

If the component measures are combined at the aggregate level, do not include the individual measure specifications below.

2a.1 Composite Numerator Statement: The sum of all successes for acute myocardial infarction process-of-care indicators, weighted by one-half the reciprocal of the share of opportunities represented by acute myocardial infarction process-of-care indicators in total opportunities, plus the sum of all successes for acute myocardial infarction outcome-of-care indicators, weighted by one-half the reciprocal of the share of opportunities represented by acute myocardial infarction outcome-of-care indicators in total opportunities.

2a.2 Numerator Time Window: July 2006 - June 2009

2a.3 Numerator Details: Successes in the following acute myocardial infarction process-of-care and outcome of care indicators:

HOSPITAL PROCESS-OF-CARE INDICATORS

1. Percent of AMI Patients Given Aspirin on Arrival (NQF #0132)
2. Percent of AMI Patients Given Aspirin at Discharge (NQF #0142)
3. Percent of AMI Patients Given ACE Inhibitor or ARB for LVSD (NQF #0137)
4. Percent of AMI Patients Given Smoking Cessation Advice/Counseling (NQF #0027)
5. Percent of AMI Patients Given Beta Blocker at Discharge (NQF #0160)
6. Percent of AMI Patients Given Fibrinolytic Medication within 30 Min. of Arrival (NQF #0164)
7. Percent of AMI Patients Given PCI within 90 Min. of Arrival (NQF #0163)

HOSPITAL OUTCOME-OF-CARE INDICATORS

1. AMI 30-day Risk-Standardized Mortality (NQF #0230)
2. AMI 30-day Risk-Standardized Readmission (NQF #0505)

2a.4 Composite Denominator Statement: The total number of opportunities for success on all acute myocardial infarction indicators used in the composite.

Comment [KP7]: 2a. The composite measure is well defined and precisely specified so that it can be implemented consistently within and across organizations and allow for comparability. Composite specifications include methods for standardizing scales across component scores, scoring rules (i.e., how the component scores are combined or aggregated), weighting rules (i.e., whether all component scores are given equal or differential weighting when combined into the composite), handling of missing data, and required sample sizes.

2a.5 **Target Population Gender** ☒ Female ☒ Male
2a.6 **Target Population Age range** Aged 18 and over.

2a.7 **Denominator Time Window:** July 2006 - June 2009.

2a.8 **Denominator Details:** Counts of process-of-care opportunities are based on hospital acute myocardial infarction quality reports. Counts of outcome-of-care opportunities are based on claims data.

2a.9 **Composite Denominator Exclusions:** Hospitals missing three or more acute myocardial infarction process-of-care indicators and one or more outcome-of-care indicator were excluded.

2a.10 **Denominator Exclusion Details:** Hospitals missing three or more of the acute myocardial infarction process-of-care indicators and one or more of the outcome-of-care indicators listed below were excluded from the composite calculation.

HOSPITAL PROCESS-OF-CARE INDICATORS

1. Percent of AMI Patients Given Aspirin on Arrival (NQF #0132)
2. Percent of AMI Patients Given Aspirin at Discharge (NQF #0142)
3. Percent of AMI Patients Given ACE Inhibitor or ARB for LVSD (NQF #0137)
4. Percent of AMI Patients Given Smoking Cessation Advice/Counseling (NQF #0027)
5. Percent of AMI Patients Given Beta Blocker at Discharge (NQF #0160)
6. Percent of AMI Patients Given Fibrinolytic Medication within 30 Min. of Arrival (NQF #0164)
7. Percent of AMI Patients Given PCI within 90 Min. of Arrival (NQF #0163)

HOSPITAL OUTCOME-OF-CARE INDICATORS

1. AMI 30-day Risk-Standardized Mortality (NQF #0230)
2. AMI 30-day Risk-Standardized Readmission (NQF #0505)

2a.11 **Stratification Details/Variables (***All information required to stratify the measure including the stratification variables, all codes, logic, and definitions***):**
None.

2a.18 **Type of Score:** Weighted score/comosite/scale 2a.19 If "Other", please describe: N/A

2a.20 **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

2a.42 **Method of Scoring/Aggregation:** other 2a.43 **If "other" scoring method, describe:**

The composite measure was calculated with a method that we have termed "Absolute Scoring Index with Reliability Weighting" (ASI-RW). The composite is actually derived by combining two sub-composites, one incorporating the process-of-care indicators and the other incorporating the outcome-of-care indicators.

The process-of-care sub-composite is derived by applying reliability weights to each individual process-of-care indicator, such that each hospital-specific indicator is based on the actual reported data for that indicator as well as the national mean for that indicator. The resulting adjusted rates are then weighted and added together to form the process-of-care sub-composite. The weight used to combine indicators is based on the national number of patients included in the indicator (denominator weighting), so that if one indicator is relevant to twice as many patients as another, the weight of that indicator in the composite is twice as large as the weight of the other. Many composite measures that NQF has approved use this patient measure opportunity basis; it has the advantage of focusing the outcome of the measurement process on the places where opportunities to provide appropriate evidence-based process care are greatest.

To calculate the outcome-of-care sub-composite, we first subtracted the individual indicators from 100 to create two desired outcomes: 1) survival rates, which replaced morality and 2) absence of readmission, which replaced readmission. We then applied denominator weighting, once again, to estimate the outcome-

of-care sub-composite.

After generating process- and outcome-of-care sub-composite scores, each is scaled by subtracting the overall domain mean and dividing by the standard deviation (a statistical process to derive a standardized score). The two sub-composites are combined using a simple average. To map the standardized composite score to a scale between zero to one hundred, we then add the lowest possible score a hospital can receive (i.e., a hospital scores zero percent on all process- and outcome-of-care indicators) and divide by the range of potential hospital scores (i.e., the difference between the highest possible score a hospital can receive, which is if a hospital scores 100 percent on all process- and outcome-of-care indicators, and the lowest possible score).

2a.44 Missing Component Scores (*Indicate how missing component scores are handled*):

The AMI composite measure is generated for all hospitals that reported data for at least four of the seven process-of-care indicators and one of the two outcome-of-care indicators. For hospitals that meet these criteria but are missing data for some of the component measures, missing values are imputed using the national mean.

2a.45 Weighting: ☐ **Equal** ☒ **Differential** **2a.46 If differential weighting, describe:**

Consistent with the approach used for the AHRQ measures, CMS used denominator weighting in constructing the process-of-care sub-composite. Denominator weighting places relatively more weight on measures that apply to relatively more patients nationally. Please see Appendix A for complete details on weighting methodology.

2a.21 Calculation Algorithm (*Describe the calculation of the measure as a flowchart or series of steps*)**:**
Please note: Complete information on the calculation algorithm, including equations, are contained in Appendix A. The text summary follows below.

STEP 1
Hospital process-of-care indicators for AMI, with a data collection period of July 2008 to June 2009, and outcome-of-care indicators for AMI, with a data collection period of July 2006 to June 2009, that are publicly reported on Hospital Compare (http://www.hospitalcompare.hhs.gov/), are combined into a single data set using hospital provider identification numbers.

STEP 2
Process-of-care indicators are reliability-weight adjusted. That is, the value of each process-of-care indicator is set equal to the weighted average of the hospital's own mean for the indicator and the national mean for the indicator. The weights are based on the between-hospital variance and the within-hospital variance in indicator scores (for more information on this adjustment, see the "Estimation of the Reliability-Weight-Adjusted Measures," which follows).(ii)

STEP 3
Hospitals missing process- or outcome-of-care indicators are imputed with the national mean.(iii) The national mean of the process-of-care indicators are estimated as a simple average of the indicators. The national mean of the outcome-of-care indicators are provided by Hospital Compare.

STEP 4
The process-of-care sub-composite score is computed using denominator weights, where the denominator weight is based on the number of hospital cases for each process-of-care indicator (see Appendix A, "Estimation of the Absolute Score Index with Reliability Weighting Composite Measure," eq. 2a.21.5).

STEP 5
The outcome-of-care sub-composite score is also computed using denominator weights (see Appendix A, "Estimation of the Absolute Score Index with Reliability Weighting Composite Measure," Equation 2a.21.6).

STEP 6
To standardize the process- and outcome-of-care sub-composite measures, each are scaled by subtracting

the overall sub-composite mean and dividing by the standard deviation. Then the average of the process- and outcome-of-care sub-composites is estimated (see Appendix A, "Estimation of the Absolute Score Index with Reliability Weighting Composite Measure," Equation 2a.21.8).

STEP 7
Lastly, in order to have a composite measure with values between zero and 100, we add the lowest possible score a hospital can receive (i.e., a hospital scores zero percent on all process- and outcome-of-care indicators) and divide by the range of potential hospital scores (i.e., the difference between the highest possible score a hospital can receive, which is if a hospital scores 100 percent on all process- and outcome-of-care indicators, and the lowest possible score) (see Appendix A, "Estimation of the Absolute Score Index with Reliability Weighting Composite Measure," Equation 2a.21.7).

ESTIMATION OF RELIABILITY-WEIGHT-ADJUSTED MEASURES

For each process-of-care indicator, the reliability-weight-adjusted indicator is equal to a weighted average of the hospital's own measure and the national mean value of the measure. In each case, the weight is a measure of the precision with which a hospital's measure has been estimated. This weighted average has been shown to be more accurate, on average, than using each hospital's individual value for the measure.

The weight is made up of two parts—the variability of the measure within each hospital, termed the "within variance" or "noise variance," and the variability across hospitals, known as the "signal variance." The weight attached to each hospital's own value for process measure k is equal to the ratio of the signal variance to the sum of the signal variance and the noise variance. As the number of observations for a hospital (njk) increases, the weight approaches one. Please see Appendix A for complete calculation details.

ESTIMATION OF THE ABSOLUTE SCORE INDEX WITH RELIABILITY WEIGHTING (ASI-RW) COMPOSITE MEASURE

We estimate the composite measure using an approach that we have termed absolute score index with reliability weighting (ASI-RW). To compute the ASI-RW, we first computed process- and outcome-of-care sub-composite scores. Using process-of-care indicators that are set equal to the weighted average of the hospital's own mean for the indicator and the national mean for the indicator (that is, reliability-weight adjusted), the process-of-care sub-composite score is computed as a denominator-weighted average of the process-of-care indicators. That is, weights of each process-of-care indicator are based on the opportunities for providing a specific recommended treatment and greater weights are placed on measures that apply to relatively more patients nationally. Similarly, the outcome-of-care sub-composite score is also estimated as a denominator-weighted average of the outcome-of-care indicators, which are reported on Hospital Compare and are risk-adjusted.

To standardize each measure, the process- and outcome-of-care sub-composite scores are scaled by subtracting the overall sub-composite mean and dividing by the standard deviation. The ASI-RW composite measure is computed using two steps. First, the average of the process- and outcome-of-care sub-composites is estimated. Then, to map the standardized composite score to a scale between zero and 100, we add the lowest possible score a hospital can receive (i.e., a hospital scores zero percent on all process- and outcome-of-care indicators) and divide by the range of potential hospital scores (i.e., the difference between the highest possible score a hospital can receive, which is if a hospital scores 100 percent on all process- and outcome-of-care indicators, and the lowest possible score). Please see Appendix A for complete calculation details.

Footnotes
ii. Hospital outcome-of-care indicators are not reliability-weight adjusted because they have been risk-standardized using a method that accounts for reliability previously, before public reporting on Hospital Compare.
iii. The use of the national mean is consistent with the approach used for the AHRQ quality composites. It is simple, already in use, and perceived as fair by providers.

2a.22 Describe the method for discriminating performance *(e.g., significance testing)*:
Please note: Complete information on the method for discriminating performance, including equations, are contained in Appendix A. The text summary follows below.

To examine meaningful differences in composite measures among hospitals, for the purpose of internal

analysis, we compared hospitals' confidence interval estimates with the overall mean and assigned hospitals into one of three performance categories: better than hospitals, if the interval estimate is entirely above the mean; no different than hospitals, if the interval estimate includes the mean; and worse than hospitals, if the interval estimate is entirely below the mean. These categories were used for illustrative analyses only and should not be assumed to be the manner in which these composites will be publicly reported.

The hospital-specific standard error is estimated by computing the variance of the composite measure and computing a square root of the variance. After we derive the standard errors for each hospital, we estimate an interval estimate around each hospital's mean composite measure. The interval estimate is a range of probable values for the composite measure that characterizes the amount of uncertainty associated with the estimate. We apply a 95 percent interval estimate, which indicates a 95 percent confidence level that the true composite measure is between the lower and upper limits of the interval.

2a.23 Sampling (Survey) Methodology *If measure is based on a sample (or survey), provide instructions for obtaining the sample (or conducting the survey) and guidance on minimum sample size (response rate)*:
N/A

2a.24 Data Source *Check all the source(s) used in the component measures.*

- ☐ Documentation of original self-assessment *(e.g., SF-36)*
- ☒ Electronic administrative data/ claims
- ☐ Electronic Clinical Data *(e.g., MDS)*
- ☐ Electronic Health/Medical Record
- ☐ External audit
- ☐ Lab data
- ☐ **M**anagement data
- ☐ Organizational policies and procedures
- ☐ Paper Medical Record/flowsheet
- ☐ Pharmacy data
- ☐ Public health data/vital statistics
- ☐ Registry data
- ☐ Survey-patient *(e.g., CAHPS)*
- ☐ Survey-provider
- ☐ Special or unique data, specify:

2a.25 Data source or collection instrument (*Identify the specific data source or data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.*):

The composite is constructed from component measures posted on the Hospital Compare website.

2a.26 Data source/data collection instrument attached ☐ **OR 2a.27 at web page URL:**
http://www.hospitalcompare.hhs.gov/

2a.29 Data dictionary/code table attached ☐ **OR 2a.30 at web page URL:**
http://www.hospitalcompare.hhs.gov/

2a.32 Level of Measurement/Analysis (*Check the level for which the measure is specified and tested*)

Clinicians: ☐ Individual ☐ Group ☐ Other
- ☒ Facility/Agency (e.g., hospital, nursing home)
- ☐ Health plan
- ☐ Integrated delivery system
- ☐ **M**ulti-site/corporate chain

Population: ☐ National ☐ Regional/network ☐ State ☐ Counties/Cities

- ☐ Prescription drug plan

Program: ☐ Disease management ☐ QIO ☐ Other

- ☐ Measured at all levels
- ☐ Other (*Please describe*):

2a.26 Care Settings (*Check the settings for which the measure is specified and tested; check all that apply*)
Ambulatory Care: ☐ Amb Surgery Center ☐ Office ☐ Clinic ☐ Emergency Dept ☐ Hospital Outpatient

- ☐ Assisted Living
- ☐ Behavioral health/psychiatric unit
- ☐ Dialysis Facility
- ☐ Emergency medical services/ambulance
- ☐ Group Home
- ☐ Home
- ☐ Hospice
- ☒ Hospital
- ☐ Long term acute care hospital
- ☐ Nursing home/ Skilled Nursing Facility (SNF)
- ☐ Rehabilitation Facility
- ☐ All settings
- ☐ Unspecified or "not applicable"
- ☐ Other (*Please describe*):

2a.38 Clinical Services (*Healthcare services being measured; all that apply.*)

Behavioral Health:
- ☐ Mental health
- ☐ Substance use treatment
- ☐ Other

Clinicians:
- ☐ Audiologist
- ☐ Chiropractor
- ☐ Dentist/Oral surgeon
- ☐ Dietician/Nutritional professional
- ☒ Nurses
- ☐ Optometrist
- ☐ PA/NP/Advanced Practice Nurse
- ☐ Pharmacist
- ☒ Physicians (MD/DO)
- ☐ Podiatrist
- ☐ Psychologist/LCSW
- ☐ PT/OT/Speech
- ☐ Respiratory Therapy
- ☐ Other

- ☐ Dialysis
- ☐ Home health
- ☐ Hospice/Palliative care
- ☐ Imaging services
- ☐ Laboratory
- ☐ Other

If the component measures are combined at the patient level and include outcomes, complete the following

2a.12 Risk Adjustment Type: ☐ No risk adjustment necessary ☐ analysis by subgroup ☐ case-mix adjustment ☐ paired data at patient level ☐ risk-adjustment devised specifically for this measure/condition ☐ risk adjustment method widely or commercially available ☐ Other (specify) **2a.13**

2a.14 Risk Adjustment Methodology/Variables (*List risk adjustment variables and describe conceptual models, statistical models, or other aspects of model or method***):**

2a.15 Detailed risk model attached ☐ **OR 2a.16 at web page URL:**

TESTING/ANALYSIS

2i. Component item/measure analysis to justify inclusion in composite

2i.1 Data/sample:

As noted in Section 1d, the purpose of the proposed composite is to summarize the process- and outcome-of-care indicators associated with treatment of AMI that are now reported under the Hospital Inpatient Quality Reporting Program. Because we do not justify the composite in terms of the behavior of those indicators, our analysis aims to document the strength of associations among them; we are interested in the extent to which our formative measure does in fact represent a single construct of well-coordinated, high-quality care.

The analysis reported here relies on data that are publicly reported on Hospital Compare. We merged process-of-care indicators for AMI collected between July 2008 and June 2009 and outcome-of-care indicators for AMI collected between July 2006 and June 2009. A total of 4,990 hospitals were reported on Hospital Compare during this time period. Of these, we estimated AMI composite measures for 2,738 hospitals, with non-mising data for at least four of the seven process-of-care indicators and at least one of the two outcome-of-care indicators.

The seven AMI hospital process-of-care indicators used in the construction of composites were drawn from Medicare hospital administrative claims data and medical record documents with discharge dates between July 2008 and June 2009. The hospital outcome-of-care indicators for 30-day risk-adjusted mortality and readmission for AMI were based on Medicare claims for hospital stays with discharge dates between July 2006 and June 2009. It is important to bear in mind that process-of-care indicators were reported for all patients and that outcome-of-care indicators were computed from claims data for Medicare patients only.

2i.2 Analytic Method:

We carried out two analyses to explore the structure of the AMI indicators. First, we examined correlations among all process- and outcome-of-care indicators. Second, we conducted an exploratory factor analysis on the same process- and outcome-of-care indicators. Results appear in Appendix A, Tables 2i.3.1 and 2i.3.2

2i.3 Results:

2i
C ☐
P ☐
M ☐
N ☐

Comment [KP8]: 2i. Component item/measure analysis (e.g., various correlation analyses such as internal consistency reliability), demonstrates that the included component items/measures fit the conceptual construct;
OR
justification and results for alternative analyses are provided.

Please see Appendix A for complete details on results. The text summary follows below.

All correlations are positive, as Table 2i.3.1 (see Appendix A) shows, though many are weak, with values below 0.10. The two time-sensitive indicators (AMI 7A and AMI 8A) exhibit low correlation with other indicators. This is probably due to the high frequency of missing values for these two measures and their replacement with the overall mean. Correlations between process- and outcome-of-care indicators are low, though consistently positive. In addition, the Cronbach's alpha is 0.48, which is slightly below the commonly desired value of 0.70.

The factor analysis of component measures (Table 2i.3.2, see Appendix A) produced a single factor with an eigenvalue greater than one. The eigenvalue for the first factor was more than 10 times that of the second factor, strongly suggesting that the component indicators represent a single underlying construct.

2j. Component item/measure analysis of contribution to variability in composite score

Comment [KP9]: 2j. Component item/measure analysis demonstrates that the included components contribute to the variation in the overall composite score;
OR
if not, justification for inclusion is provided.

2j.1 Data/sample:

As noted in Section 1d, the purpose of the proposed composite is to summarize the process- and outcome-of-care indicators associated with treatment of AMI that are now reported under the Hospital Inpatient Quality Reporting Program. Because we do not justify the composite in terms of the behavior of individual indicators, our analysis aims to document their contributions to the measure.

Analysis of the contribution of component items to the variability in composite scores uses data that are publicly reported on Hospital Compare. We merged process-of-care indicators for AMI with a data collection period of July 2008 to June 2009 and outcome-of-care indicators for AMI with a data collection period of July 2006 to June 2009. A total of 4,990 hospitals were reported on Hospital Compare during this time period. Of these, we estimated composite measures for 2,738 hospitals, for which less than or equal to three process-of-care indicators and less than or equal to one outcome-of-care indicator is missing.

The seven hospital process-of-care indicators related to AMI that are used in the construction of the AMI composite are calculated from Medicare hospital administrative claims data and medical record documents with discharge dates between July 2008 and June 2009. The hospital outcome-of-care indicators for 30-day risk-adjusted mortality and readmission for AMI are based on Medicare claims for hospital stays with discharge dates between July 2006 and June 2009.

2j.2 Analytic Method:

We compare the percentage change in (1) the variance and (2) the inter-quartile range (IQR) of the process- and outcome-of-care sub-composites when a process- or outcome-of-care indicator is removed before normalization. Results appear in Appendix A, Table 2j.3.1.

2j.3 Results:

Please see Appendix A for complete details on results. The text summary follows below.

In Table 2j.3.1 (Appendix A), the positive values indicate that addition of the component indicator tends to reduce the variance or IQR. Only one indicator, AMI2 (aspirin at discharge), exhibits a nontrivial positive effect on the composite variance, probably because of its relatively strong positive correlation with other component indicators (see Table 2i.3.1). Because the outcome domain contains only two component indicators, readmission and mortality both have strong negative effects on the variance of the sub-composite measure. The strong variance-reducing effect of readmission appears to be the result of its tight distribution (see Table 2l.3.2, Appendix A).

2j
C ☐
P ☐
M ☐
N ☐

2k. Analysis to support differential weighting of component scores

Comment [KP10]: 2k. The scoring/aggregation and weighting rules are consistent with the conceptual construct. (Simple, equal weighting is often preferred unless differential weighting is justified. Differential weights are determined by empirical analyses or a systematic assessment of expert opinion or values-based priorities.)

2k.1 Data/sample:

In constructing the composite, individual component indicators are weighted, in each instance, by the number of observations for the indicator. The most frequently reported indicators therefore affect the composite most strongly. In addition, the weighting scheme tends to reduce the variance of the composite,

2k
C ☐
P ☐
M ☐
N ☐

though this effect might be muted if individual indicators have similar distributions.

Testing to support differential weighting of composite scores uses data that are publicly reported on Hospital Compare by CMS. We merged process-of-care indicators for AMI with a data collection period of July 2008 to June 2009 and outcome-of-care indicators for AMI with a data collection period of July 2006 to June 2009. A total of 4,990 hospitals were reported on Hospital Compare during this period. Of these, we estimated AMI composite measures for 2,738 hospitals, for which less than or equal to three process-of-care indicators and less than or equal to one outcome-of-care indicator is missing.

The seven hospital process-of-care indicators related to AMI that are used in the construction of composites are drawn from Medicare hospital administrative claims data and medical record documents with discharge dates between July 2008 and June 2009. The hospital outcome-of-care indicators for 30-day risk-adjusted mortality and readmission for AMI are based on Medicare claims for hospital stays with discharge dates between July 2006 and June 2009.

2k.2 **Analytic Method:**

We compare the distribution of the AMI composite measure with equal and differential weighting. Please see Appendix A for complete details on the analytic method, including equations.

2k.3 **Results:**

Please see Appendix A for complete details on results. The text summary follows below.

Table 2k.3.1 (Appendix A) displays the distribution of the AMI composite measure with equal and differential weighting. As the table shows, denominator weighting has little effect on the distribution of the composite. The median is slightly larger when denominator weighting is used, and the inter-quartile range is somewhat smaller.

2k.4 **Describe how the method of scoring/aggregation achieves the stated purpose and represents the quality construct:**

The objective of the composite is to summarize the component measures in a useful and scientifically acceptable manner.

Because composites are most useful to consumers if differences in composite values are clinically and statistically meaningful and reflect true differences in underlying quality, CMS entered component measures as values, not ranks, and adjusted those values for reliability. CMS entered component measures as values rather than ranks to prevent slight differences in composite values from producing large differences in composite values, as can occur when indicators are tightly distributed across hospitals. CMS also adjusted the component indicators for reliability so that random variation did not drive small hospitals to extremes; 30-day outcome measures are adjusted for reliability before publication on Hospital Compare. Process measures are not adjusted for reliability before publication; the adjustment is made as part of the compositing process.

In addition, because composites are more useful to consumers if they emphasize measures that are relevant to a larger number of consumers, CMS constructed the process- and outcome-of-care composite scores using weights based on national denominators.

When sample sizes are equal, each component process measure contributes equally to the AMI process-of-care domain score. The same is true for each component outcome-of-care indicator. Thus a hospital that improves in any component will necessarily produce an increase in its composite score. Hospitals can therefore choose where to focus improvement efforts in evidence-based processes of care. Similar logic applies to the outcome-of-care domain score. The composite thus fully reflects the AMI process and outcome-of-care indicators and represents the quality construct expressed earlier.

2k.5 **Indicate if any alternative scoring/aggregation methods were tested and why not chosen:**

In addition to the preferred compositing approach, ASI-RW, two alternative scoring methods were analyzed.

These are referred to as (1) the absolute scoring index (ASI) and (2) the modified relative quality index (MRQI).

1. Absolute Scoring Index
ASI is similar to CMS' preferred approach but component indicators are not reliability-adjusted. To compute the ASI composite measure, process- and outcome-of-care sub-composite scores are first computed as the equally-weighted average of the indicators. These process and outcome scores are then scaled by subtracting the overall domain mean and dividing by the standard deviation. The ASI is then computed as the average of these two scaled means. The measure is then mapped to a scale between zero and 100, for ease of interpretation.

2. Modified Relative Quality Index
MRQI is similar to CMS' preferred approach but component indicators are not reliability-adjusted and enter the composite as ranks, not values. To compute the MRQI composite measure, scores for process- and outcome-of-care sub-composites are computed as the mean of the hospital's ranks for each indicator. The composite score is then computed as the simple mean of the two domain scores. It is closely related to the relative quality index (RQI) as described by Tompkins et al. (1) (January 2009, August 2009).

In Table 2k.5.1 (see Appendix A), we present distributions of the three alternative scoring methods. Broadly speaking, the distributions for ASI-RW, the preferred approach, look quite similar to the distribution for ASI.(iv) The difference is that the reliability adjustment has reduced the likelihood of erroneously classifying small hospitals as outliers due to random variation in measured performance by pulling them toward the mean of the distribution, though this is not visible in the table itself.

Results for MRQI show a more balanced distribution with medians close to means and less pronounced clustering in the upper half of the distribution, although there is still some clustering. Note that although this approach makes the distribution look more balanced, it does not address the fundamental problems of highly clustered performance on the underlying measures, small numbers of observations, and difficulty identifying meaningful differences in performance.

Citations
1. Tompkins, Chris, Grant Ritter, Andrew Ryan, Wato Nsa, James Burgess, and Dale Bratzler. Composite Measures for Public Reporting: Final Report. January 16, 2009 (updated August, 2009).

Footnotes
iv. Although the shrinking process pulls the process-of-care indicators toward the mean, shrinking does not result in a smaller standard deviation of the distribution of the final composite values because each domain is normalized.

2l. Analysis of missing component scores

Comment [KP11]: 2l. Analysis of missing component scores supports the specifications for scoring/aggregation and handling of missing component scores.

2l.1 **Data/sample:**

The seven hospital process-of-care indicators used in the construction of the composite are drawn from reports for patients discharged between July 2008 and June 2009. Outcome-of-care indicators for 30-day risk-adjusted mortality and readmission for AMI are based on Medicare claims for stays with discharge dates between July 2006 and June 2009.(v) Process and outcome indicators were reported on Hospital Compare for 4,990 hospitals during this period.

Because some hospitals did not report all indicators, some method for dealing with missing data was required. In order to compute a composite AMI measure for the greatest possible number of hospitals, we followed two principles:

1. All seven process-of-care indicators and both outcome-of-care indicators were included.

2. A composite measure was computed for each hospital that reported four or more process-of-care indicators and at least one outcome-of-care indicator.

AMI composites were computed for all hospitals that satisfied the second item. The national mean was used to impute a value for any missing process- or outcome-of-care indicators. AMI composite measures were

2l
C☐
P☐
M☐
N☐

computed for 2,738 hospitals.

Footnotes
v. The reporting periods for the process- and outcome-of-care measures represent the most recent version of both measures available on the Hospital Compare website, at the time of this NQF submission.

2l.2 **Analytic Method:**

We used two approaches to conduct analysis of missing component scores. First, we tested to identify if there were differences in the composite measure for hospitals that did not fit the criteria stated previously. That is, we assessed whether composite measures differed significantly when our sample included composites with data for at least four process-of-care indicators and at least one outcome-of-care indicator (2,738 hospitals in total), compared with hospitals that did not have data for at least four process-of-care indicators and at least one outcome-of-care indicator (2,252 hospitals). Distributions of hospital composite scores were compared by the number of hospitals missing process- and outcome-of-care indicators. Second, we compared (1) distributions of process- and outcome-of-care indicators; (2) distributions of composite measures; (3) Spearman (rank) correlations; and (4) kappa statistics for hospital quartiles, with and without imputation of the national mean. Results appear in Appendix A, Tables 2l.3.1 to 2l.3.4.

2l.3 **Results:**

Please see Appendix A for complete details on results. The text summary follows below.

Hospitals are more likely to fail to meet the required minimum number of outcome-of-care indicators (one) than to fail to meet the minimum number of process-of-care indicators (four). That is, four or more process-of-care indicators are missing for 1,783 hospitals (35.7%), while both outcome-of-care indicators are missing for 2,102 hospitals (42.1%). Note that all component measures are missing for 1,190 hospitals (23.9%).

The distributions of component measures, shown in Table 2l.3.2 (Appendix A), are largely the same whether or not missing values are imputed. The clear exception is the distribution of process-of-care indicators AMI 7A and AMI 8A (fibrinolytic medication within 30 minutes of arrival and PCI within 90 minutes of arrival, respectively). We imputed these two measures far more often than other measures.

Table 2l.3.3 (Appendix A) shows the distribution of composites for the 2,738 hospitals for which at least four process measures and one outcome measure are reported. Composites with no imputation simply drop missing component measures from the calculation; composites with imputation use the national mean in place of the missing measure. As the table shows, the two different procedures yield nearly identical distributions for the AMI composite.

Table 2l.3.4 (Appendix A) shows the association between imputed and non-imputed measures by quartile. More than three-fourths of hospitals lie on the diagonal, occupying the same quartile for composite values using imputed and non-imputed component measures. The Spearman correlation coefficient and the kappa statistic both indicate a strong positive relationship between the two.

2b. Reliability testing of composite score

2b.1 **Data/sample** *(description of data/sample and size)*:

The reliability of the proposed AMI composite measure is informed by the reliability of the component scores on which it is based. Two reports, one by Williams et al. (2006) and the other by the Government Accountability Office (GAO) (2006), provide insight into component measure reliability:

Williams, SC, Watt, A, Schmaltz, SP, Koss, RG, & Loeb, JM. Assessing the reliability of standardized performance indicators. Int J Qual Health Care. 2006 Jun;18(3):246-55. Epub 2006 Jan 23.

- Williams et al. (2006) examined the reliability of all seven AMI process-of-care indicators that make up the AMI composite. Their sample included 30 hospitals, representing a diverse range of geographic locations, sizes, settings (urban/rural), and ownership categories (profit/not-for-profit); 19 of these collected AMI data. A randomly selected set of deidentified, previously abstracted medical records was

2b
C☐
P☐
M☐
N☐

Comment [KP12]: 2b. Reliability testing of the composite measure demonstrates the results are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period.

transmitted from the hospitals' performance measurement vendors and AMI process-of-care indicators were reabstracted following guidelines from the Specification Manual for National Implementation of Hospital Core Measures. Sample sizes used to calculate each measure generally ranged from 100-200 cases, though for AMI-4 (smoking cessation counseling) and AMI-8A (first PCI time) the sample size was fewer than 50.

United States. Government Accountability Office. Report to the Committee on Finance, U.S. Senate. Hospital Quality Data: CMS Needs More Rigorous Methods to Ensure Reliability of Publicly Released Data. Report No. GAO-06-54, Jan. 31, 2006

- The 2006 GAO report summarizes CMS' process to assess the reliability of the measures currently reported on Hospital Compare and reports the results of this process for hospital discharges between January 1, 2004, and June 30, 2004. CMS' contractor, CDAC (Clinical Data Abstraction Center), assesses the reliability of the component measures on a quarterly basis. This assessment uses a sample of five randomly selected patient records from each hospital participating in the Hospital Inpatient Quality Reporting Program, which includes hospitals from all states but Maryland and Puerto Rico.(vi)

Footnotes
vi. As a result of the GAO report, in 2010 this process changed so that CDAC instead reviews 12 patient records from a randomly selected sample of 800 hospitals.

2b.2 Analytic Method *(type of reliability & rationale, method for testing)*:

Williams, SC, Watt, A, Schmaltz, SP, Koss, RG, & Loeb, JM. Assessing the reliability of standardized performance indicators. Int J Qual Health Care. 2006 Jun;18(3):246-55. Epub 2006 Jan 23.

- Reliability was assessed using percentage agreement for continuous variable elements and chance-corrected agreement using Cohen's kappa for binary data elements.

United States. Government Accountability Office. Report to the Committee on Finance, U.S. Senate. Hospital Quality Data: CMS Needs More Rigorous Methods to Ensure Reliability of Publicly Released Data. Report No. GAO-06-54, Jan. 31, 2006

- For each hospital, data are deemed reliable if there is 80 percent or greater agreement between the hospital quality data previously submitted to CMS and the CDAC reabstraction results.

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

Williams, SC, Watt, A, Schmaltz, SP, Koss, RG, & Loeb, JM. Assessing the reliability of standardized performance indicators. Int J Qual Health Care. 2006 Jun;18(3):246-55. Epub 2006 Jan 23.

- Table 2b.3.1 (Appendix A) summarizes the reliability statistics for the AMI measures that are included in the proposed composite. Using the standards proposed by Landis & Koch (1977),(1) the resulting kappas indicate almost perfect agreement (kappa > 0.81) for three of the measures, substantial agreement (kappa ranging from 0.61 to 0.80) for one measure, and moderate agreement (kappa ranging from 0.41 to 0.60) for two measures. Although a kappa was not calculated for AMI-8A (first PCI time), the authors report 64.7 percent agreement for this measure.

United States Government Accountability Office. Report to the Committee on Finance, U.S. Senate. Hospital Quality Data: CMS Needs More Rigorous Methods to Ensure Reliability of Publicly Released Data. Report No. GAO-06-54, Jan. 31, 2006

- The GAO report, which looked at reporting from January 1, 2004, through June 30, 2004, found that 90 percent of hospitals exceeded the 80 percent reliability threshold.

Citations
1. Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. Biometrics 33: 159-174

2c. Validity testing of composite score

2c.1 Data/sample *(description of data/sample and size)*:

The testing of the validity of the component scores uses two sets of data. The first data set merges process- and outcome-of-care indicators for AMI with a data collection period of July 2008 to June 2009. The second data set merges process- and outcome-of-care indicators for AMI with a data collection period of July 2007 to June 2008. Composite measures are calculated from these two separate periods and compared, with the assumption that a valid composite measure should show minimal change on a year-to-year basis.

Across these two data collection periods, 1,747 hospitals had valid composite measures for AMI.

2c.2 Analytic Method *(type of validity & rationale, method for testing)*:

Using the two sets of data, we compared composite measures across the two years using the Spearman (rank) correlation coefficient to evaluate the predictive validity of the composite measure over time.

2c.3 Testing Results *(statistical results, assessment of adequacy in the context of norms for the test conducted)*:

Please see Appendix A for complete details on results. The text summary follows below.

The Spearman correlation between composite measures computed in 2007–2008 and 2008–2009 was 0.43 ($p < 0.001$), indicating moderate predictive validity of the composite. (See Appendix A, Table 2c.3.1.) A large number of hospitals (about 40 percent) lie on the diagonal, such that the same hospital quartiles for composite values were occupied during 2007–2008 and 2008–2009. In contrast, very few hospitals (about 5 percent) occupy the first quartile in 2007–2008 and the fourth quartile in 2008–2009, and vice versa. Across the two separate periods, about 40 percent of hospitals' categorizations differ by one quartile (that is, during 2008–2009, a hospital was one quartile above or below its categorization in 2007–2008). This discrepancy appears to be a result of the tight distribution of the process- and outcome-of-care indicators, as shown in Table 2l.3.2 (Appendix A).

2c
C☐
P☐
M☐
N☐

Comment [KP13]: 2c. Validity testing of the composite measure demonstrates that the measure reflects the quality of care provided, adequately distinguishing good and poor quality. If face validity is the only validity addressed, it is systematically assessed.

2f. Identification of Meaningful Differences in Performance Across Entities

2f.1 Data/sample from Testing or Current Use *(description of data/sample and size)*:

Testing to identify meaningful differences in performance of composite scores uses data that are publicly reported on Hospital Compare by CMS. We merged process-of-care indicators for AMI with a data collection period of July 2008 to June 2009 and outcome-of-care indicators for AMI with a data collection period of July 2006 to June 2009. A total of 4,990 hospitals were reported on Hospital Compare during this period. Of these hospitals, we estimated composite measures for 2,738, for which less than or equal to three process-of-care indicators and less than or equal to one outcome-of-care indicator is missing.

The seven hospital process-of-care indicators related to AMI that are used in the construction of composites are drawn from Medicare hospital administrative claims data and medical record documents with discharge dates between July 2008 and June 2009. The hospital outcome-of-care indicators for 30-day risk-adjusted mortality and readmission for AMI are based on Medicare claims for hospital stays with discharge dates between July 2006 and June 2009.

2f.2 Methods to identify statistically significant and practically/meaningfully differences in performance *(type of analysis & rationale)*:

To examine meaningful differences in composite measures across hospitals, we compared hospitals' confidence interval estimates with the overall mean and assigned hospitals into one of three performance categories: better than hospitals, if the interval estimate is entirely above the mean; no different than hospitals, if the interval estimate includes the mean; and worse than hospitals, if the interval estimate is entirely below the mean. These performance categories do not reflect how the composites will ultimately be displayed on Hospital Compare.

2f
C☐
P☐
M☐
N☐

Comment [KP14]: 2f. Methods for scoring and analysis of the composite measure allow for identification of statistically significant and practically/ clinically meaningful differences in performance.

2f.3 **Provide Measure Scores from Testing or Current Use** *(description of scores, e.g., distribution by quartile, mean, median, SD, etc.; identification of statistically significant and meaningfully differences in performance)* **:** Please see Appendix A for complete results of measure scores from testing. The text summary follows below. CMS has not decided how it will ultimately display hospital performance to consumers on Hospital Compare or to providers in hospital-specific reports. Table 2f.3.1 in Appendix A provides the number of hospitals in each of the three performance categories (better/no different/worse than the mean). These performance categories do not reflect how the composites will ultimately be displayed on Hospital Compare.	
2h. Disparities in Care 2h.1 **If measure is stratified, provide stratified results** *(scores by stratified categories/cohorts)***:** The measure is not stratified. 2h.2 **If disparities have been reported/identified, but measure is not specified to detect disparities, provide follow-up plans:** No disparities have been reported/identified.	**2h** C☐ P☐ M☐ N☐ NA☐
If the component measures are combined at the patient level, complete 2d. **2d. Exclusions Justified** 2d.1 **Summary of Evidence supporting exclusion(s):** 2d.2 **Citations for Evidence:** 2d.3 **Data/sample** *(description of data/sample and size)***:** 2d.4 **Analytic Method** *(type analysis & rationale)***:** 2d.5 **Testing Results** *(e.g., frequency, variability, sensitivity analyses)***:**	**2d** H☐ M☐ L☐ N☐ NA☐
If the component measures are combined at the patient level and include outcomes, complete 2e. **2e. Risk Adjustment** 2e.1 **Data/sample** *(description of data/sample and size)***:** 2e.2 **Analytic Method** *(type of risk adjustment, analysis, & rationale)***:** 2e.3 **Testing Results** *(risk model performance metrics)***:** 2e.4 **If outcome or resource use measure is not risk adjusted, provide rationale:**	**2e** H☐ M☐ L☐ N☐ NA☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties?*	**2**
Steering Committee: Overall, to what extent was the criterion, *Scientific Acceptability of Measure Properties***, met?** **Rationale:**	*2* C☐ P☐ M☐ N☐
3. USABILITY	
Extent to which intended audiences (e.g., consumers, purchasers, providers, policy makers) can understand the results of the measure and are likely to find them useful for decision making. (composite measure evaluation criteria)	**Eval**
3a. Meaningful, Understandable, and Useful Information	**3a** C☐

Comment [KP15]: 2h. If disparities in care have been identified, measure specifications, scoring, and analysis allow for identification of disparities through stratification of results (e.g., by race, ethnicity, socioeconomic status, gender);
OR
rationale/data justifies why stratification is not necessary or not feasible.

Comment [KP16]: 2d. Clinically necessary measure exclusions are identified and must be:
- supported by evidence of sufficient frequency of occurrence so that results are distorted without the exclusion;

AND
- a clinically appropriate exception (e.g., contraindication) to eligibility for the measure focus;

AND
- precisely defined and specified:
  - if there is substantial variability in exclusions across providers, the measure is specified so that exclusions are computable and the effect on the measure is transparent (i.e., impact clearly delineated, such as number of cases excluded, exclusion rates by type of exclusion);

if patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that it strongly impacts performance on the measure and the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).

Comment [KP17]: 2e. For outcome measures and other measures (e.g., resource use) when indicated:
- an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified and is based on patient clinical factors that influence the measured outcome (but not disparities in care) and are present at start of care; OR

rationale/data support no risk adjustment.

Comment [KP18]: 3a. Demonstration that information produced by the composite measure is meaningful, understandable, and useful to the intended audience(s) for both public reporting (e.g., focus group, cognitive testing) and informing quality improvement (e.g., quality improvement initiatives).

3a.1 Current Use: ☐ In use ☒ Not in use

P☐ **M**☐ **N**☐

3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large) (*If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s). If not publicly reported, state the plans to achieve public reporting within 3 years*)**:**

Following NQF endorsement, the proposed measure will undergo a national dry run in advance of implementation on Hospital Compare. The dry run is currently slated for the second quarter of 2011. The dry run will include the following steps:

- Standard Data Processing System (SDPS) memos will be sent to hospitals and QIOs with public reporting contacts announcing the dry run.
- Confidential draft hospital-specific reports (HSRs) will be made available to hospitals via the "My QualityNet" website, with supporting materials describing the methods and handling of constituent measures.
- A mock report containing simulated data but describing methods in full will also be published on QualityNet.
- A 30-day comment period will be opened in order to receive hospital feedback.
- Nationwide webinars will be held in order to review the dry run process and the methodology used to derive the composite measures, and to explain the summary data provided in the HSRs.
- A summary report of the dry run will be produced, including implications for reporting the measures.

Following this process, public reporting is expected on Hospital Compare sometime in 2011.

3a.3 If used in other programs/initiatives (*If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s). If not used for QI, state the plans to achieve use for QI within 3 years***):**

Following NQF endorsement and a national dry run, CMS plans to report this composite publicly on Hospital Compare. CMS' current timetable calls for this public reporting to occur in 2011. CMS' experience indicates that hospitals closely scrutinize measures reported on Hospital Compare and consider these results as part of their quality improvement efforts.

Testing of Interpretability (*Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement*)

3a.4 Data/sample *(description of data/sample and size)***:**

Several studies suggest that the proposed composite measure will improve consumer understanding of hospital performance for AMI patients and will be an asset to clinicians. In work that is directly relevant to the proposed measure, Borck et al. (2009) held a series of focus groups that evaluated consumer and clinician understanding of condition-specific composite measures for AMI, HF, PN, and SCIP that are very similar to the proposed measure. Their work also evaluated understanding of AHRQ and HCAHPS composite measures. In addition, work by Smith et al. (2005) examined the interpretability of Hospital Compare data, including several of the component measures in the proposed composite. A further study by Peters et al. (2007) also provides insight into consumer understanding of publicly reported hospital quality measures; L&M Policy Research LLC specifically reports on consumer understanding of the readmissions outcome measure, one of two possible outcome-of-care indicators included in this composite.

Borck, M, Thomas, C, & Gerteis, M. Transparency in Public Reporting: Consumer Testing and Enhancements to CMS's Compare Tools: Topline Summary of Findings from Round #1 Interviews with Consumers, April 9, 2009, and Topline Summary of Findings from Round #2 Interviews with Consumers and Physicians, Composite measures of quality for Hospital Compare, June 11, 2009. Memoranda to the Centers for Medicare & Medicaid Services.

- Round 1 – Borck et al. (2009) used a convenience sample of 21 consumers in the Baltimore, Maryland, area. Participants ranged in age from 45 to 70; 67 percent were women, and 48 percent were Medicare beneficiaries.
- Round 2 – Borck et al. (2009) used a convenience sample of 18 consumers and five physicians from the

Miami, Florida, area. The group ranged in age from 45 to 70; most of the group's members were men and Medicare beneficiaries.

Smith, F, Gerteis, M, Burnes, A, Gerteis, J, Crelia, S, & Silva, N. Usability Testing of the "Hospital Compare" Website. Final Report to the Centers for Medicare & Medicaid Services. August 29, 2005.
▫Smith et al. (2005) used a sample of 51 consumers and 40 health care providers to assess their ability to understand Hospital Compare content and navigate the user interface website. Among the consumers, 47 of 51 (92%) were older than 65, and of the over-65 group, 53 percent were Medicare beneficiaries at risk for heart disease. Among the health care providers, 30 percent were nurses, 38 percent were primary care physicians, and the remainder were cardiologists and pulmonologists.

Peters, E, Dieckmann, N, Dixon, A, Hibbard, JH, & Mertz, CK. Less is more in presenting quality information to consumers. Med Care Res Rev. 2007 Apr;64(2):169-90.
▫ Peters et al. (2007) employed a convenience sample of employment-age adults (ages 18 to 64, mean age of 37, 48 percent female, and 76 percent white) to determine whether providing only the most important quality information increases comprehension and information use. Half of the sample had lower levels of education (high school or less), 45 percent had health insurance, and 74 percent had an annual household income of less than $20,000.

L&M Policy Research LLC. Report to the Centers for Medicare & Medicaid Services: Recommendations for Incorporating Hospital Readmission Data into the Hospital Compare Website. January 29, 2009.
▫This effort entailed two rounds of consumer testing, the first of which focused on general understanding of hospital readmission measures and how they are calculated, as well as the fact that the measures are for readmission within 30 days and calculated from Medicare fee-for-service data. The sample for this round included 10 adult consumers ages 50 to 70, most of whom were previously diagnosed with heart disease; 8 caregivers ages 40 to 60; and 6 physicians who were primary care physicians, cardiologists, and pulmonologists.

3a.5 Methods *(methods, e.g., focus group, survey, QI project)*:

Borck, M, Thomas, C, & Gerteis, M. Transparency in Public Reporting: Consumer Testing and Enhancements to CMS's Compare Tools: Topline Summary of Findings from Round #1 Interviews with Consumers, April 9, 2009, and Topline Summary of Findings from Round #2 Interviews with Consumers and Physicians, Composite measures of quality for Hospital Compare, June 11, 2009. Memoranda to the Centers for Medicare & Medicaid Services.
▫ Borck et al. (2009) used a mock Hospital Compare website that presented the composite quality measures of interest. Using a standard interview protocol, in-depth, one-on-one discussions assessed comprehension of composite measures, organization and presentation of the site, and composite labels and descriptions.

Smith, F, Gerteis, M, Burnes, A, Gerteis, J, Crelia, S, & Silva, N. Usability Testing of the "Hospital Compare" Website. Final Report to Centers for Medicare & Medicaid Services. August 29, 2005.
▫ Smith et al. (2005) tested consumers' and health providers' ability to understand and use the Hospital Compare website using both in-depth, one-on-one interviews and dyads (interviews that involve two respondents and one interviewer). Using a Hospital Compare website prototype, participants first navigated the website independently and then responded to a series of open-ended questions using an approved protocol during an approximately two-hour period.

Peters, E, Dieckmann, N, Dixon, A, Hibbard, JH, & Mertz, CK. Less is more in presenting quality information to consumers. Med Care Res Rev. 2007 Apr;64(2):169-90.
▫ Peters et al. (2007) assigned participants to one of three groups, each of which was presented with hospital quality data in a different format. In the first group, data on cost, quality, and nonquality information was unordered. In the second, cost and quality data was highlighted and presented first and nonquality information was presented last and not emphasized. In the final group, only cost and quality information was shown, and quality information was highlighted. Within each of these groups, respondents were then shown information about three hospitals and asked to choose a hospital and answer a series of questions.

L&M Policy Research LLC. Report to the Centers for Medicare & Medicaid Services: Recommendations for Incorporating Hospital Readmission Data into the Hospital Compare Website. January 29, 2009.
▫ Participants were shown paper-based mock-ups of hospital quality data and asked to compare hospitals and select a hospital for themselves and their family members.

3a.6 Results *(qualitative and/or quantitative results and conclusions)*:

Borck, M, Thomas, C, & Gerteis, M. Transparency in Public Reporting: Consumer Testing and Enhancements to CMS's Compare Tools: Topline Summary of Findings from Round #1 Interviews with Consumers, April 9, 2009, and Topline Summary of Findings from Round #2 Interviews with Consumers and Physicians, Composite measures of quality for Hospital Compare, June 11, 2009. Memoranda to the Centers for Medicare & Medicaid Services.
▫ This work yielded several important results that are directly relevant to the proposed condition-specific composite measure. Most significantly, all respondents from Round 1 correctly interpreted the star ratings for the condition-specific composites (AMI, HF, PN, and SCIP) and the HCAHPS composite measure. Round 1 also revealed that almost all participants preferred more descriptive definitions of the composites; specifically, those included a list of all the component measures making up the composite. Similarly, in Round 2 respondents were also able to interpret the star ratings for condition-specific quality ratings of composites and the HCAHPS composite correctly. However, some respondents in Round 2 did not understand that the condition-specific composite ratings included all of the individual component measures. These results indicate that the proposed condition-specific composite, which is very similar to the condition-specific measures evaluated by Borck et al. (2009), should also be easy for consumers to use. Moreover, any composite definition posted on Hospital Compare should include a list of all component measures.

Smith, F, Gerteis, M, Burnes, A, Gerteis, J, Crelia, S, & Silva, N. Usability Testing of the "Hospital Compare" Website. Final Report to Centers for Medicare & Medicaid Services. August 29, 2005.
▫ This early analysis of Hospital Compare's usability revealed that the amount of information available on the website tended to overwhelm consumers and that detailed information about interpretation added to this sense of overload. The provider participants concurred with this sentiment. Although these results certainly suggest certain challenges in making hospital quality data user friendly, the proposed composite measure aims to address this issue by creating a single benchmark that enables consumers to evaluate the quality of care at a given hospital for a given condition.

Peters, E, Dieckmann, N, Dixon, A, Hibbard, JH, & Mertz, CK. Less is more in presenting quality information to consumers. Med Care Res Rev. 2007 Apr;64(2):169-90.
▫ Similar to Smith et al. (2005), Peters et al. (2007) determined that less is more with regard to consumer understanding of hospital quality data. They found that consumer comprehension was highest when only the most relevant quality information was shown and highlighted relevant to the other information. Specifically, 62 percent of respondents chose the highest quality hospital Y when only the quality information was shown; in the other two formats it was selected by 48 percent (ordered group) and 40 percent (unordered group). Such results reinforce the idea that a composite measure can enhance the utility of hospital quality data for consumers.

L&M Policy Research LLC. Report to the Centers for Medicare & Medicaid Services: Recommendations for Incorporating Hospital Readmission Data into the Hospital Compare Website. January 29, 2009.
▫ This work suggests that a readmission measure is open to misinterpretation by consumers. For example, many participants in this study thought that readmission was a positive outcome because it meant that the hospital was providing follow-up care. In the proposed composite measure, discharges not followed by readmission improve the composite score. Although it is important to describe how the composite is created, this example highlights the need to define the composite in a simple, direct manner.

3b/3c. Relation to other NQF-endorsed measures
Identify similar or related NQF-endorsed measures to components and/or composite

3b.1 NQF # and Title of similar or related measures:

All components of this composite measure are NQF-endorsed. However there are currently no NQF-endorsed composite measures that provide a single indication of a hospital's quality of care for AMI patients. In that they also serve to provide a single, consumer-friendly indication of a hospital's quality of care as it relates to either patient safety or mortality for selected conditions, the proposed measure is similar in intent to the

following:

NQF #0531. Patient Safety for Selected Indicators (endorsed June 19, 2009/AHRQ)
NQF #0530. Mortality for Selected Conditions (endorsed June 19, 2009/AHRQ)

However, the proposed measure is condition-specific and intended to summarize the measures on Hospital Compare; thus, it provides unique and additive value above and beyond these measures.

(for NQF staff use) **Notes on similar/related endorsed or submitted measures:**

3b. Harmonization

3b.2 Are the component measure specifications harmonized, or if not, why?

The component measures are harmonized within each distinct domain of the composite (that is, processes of care and outcomes of care). Within the process domain, all component measures are reported as percentages; in the outcomes domain, both component measures are reported as rates.

3b C☐ P☐ M☐ N☐ NA☐

Comment [KP19]: 3b. The component measure specifications are harmonized.

3c. Distinctive or Additive Value

3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:

The proposed composite measure offers a condition-specific summary of the inpatient quality measures that CMS has adopted for its Hospital Inpatient Quality Reporting Program, related to the quality of care for AMI patients.

Comment [KP20]: 3c. Review of existing endorsed measures and measure sets demonstrates that the composite measure provides a distinctive or additive value to existing NQF-endorsed measures (e.g., provides a more complete picture of quality for a particular condition or aspect of healthcare).

5.1 Competing Measures If this measure is similar to measure(s) already endorsed by NQF (i.e., on the same topic and the same target population), describe why it is a more valid or efficient way to measure quality:

There are no currently endorsed composite measures on this topic or population.

3c C☐ P☐ M☐ N☐

Comment [k21]: 5. Demonstration that the measure is superior to competing measures - new submissions and/or endorsed measures (e.g., is a more valid or efficient way to measure).

3d. Decomposition of Composite

3d.1 Describe the information that is available from decomposing the composite into its components:

The component measures include the following information:

1. Percent of AMI Patients Given Aspirin on Arrival
2. Percent of AMI Patients Given Aspirin at Discharge
3. Percent of AMI Patients Given ACE Inhibitor or ARB for LVSD
4. Percent of AMI Patients Given Smoking Cessation Advice/Counseling
5. Percent of AMI Patients Given Beta Blocker at Discharge
6. Percent of AMI Patients Given Fibrinolytic Medication within 30 Minutes of Arrival
7. Percent of AMI Patients Given PCI within 90 Minutes of Arrival
8. Acute Myocardial Infarction (AMI) 30-day Mortality
9. Acute Myocardial Infarction (AMI) 30-day Readmission

3d C☐ P☐ M☐ N☐

Comment [KP22]: 3d. Data detail is maintained such that the composite measure can be decomposed into its components to facilitate transparency and understanding.

3e. Achieved stated purpose

3e.1 Describe how the scores from testing or use reported in 2f demonstrate that the composite achieves the stated purpose:

The scores demonstrate a range of performance on the AMI process and outcome quality measures. Testing of composite scores identified hospitals that perform significantly above and below the national mean of these scores. The scores thus reflect the underlying hospital performance regarding the quality measures for AMI, achieving the purpose of the composite.

3e C☐ P☐ M☐ N☐

Comment [KP23]: 3e. Demonstration (through pilot testing or operational data) that the composite measure achieves the stated purpose/objective.

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability*? **3**

Steering Committee: Overall, to what extent was the criterion, *Usability*, met?
Rationale:

3 C☐ P☐ M☐ N☐

4. FEASIBILITY

Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement. (composite measure evaluation criteria)

Eval

4a. Data Generated as a Byproduct of Care Processes

4a.1 **How are all the data elements that are needed to compute measure scores generated?** (*Check all that apply)*

☐ Data are generated as a byproduct of care processes during care delivery (*Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition)*

☒ Coding/abstraction performed by someone other than person obtaining original information *(e.g., DRG, ICD-9 codes on claims; chart abstraction for quality measure, registry)*

☐ Survey

☐ Other *(e.g., patient experience of care surveys, provider surveys, observation)*, *Please describe*:

4a C☐ P☐ M☐ N☐

Comment [KP24]: 4a. For clinical composite measures, overall the required data elements are routinely generated concurrent with and as a byproduct of care processes during care delivery.

4b. Electronic Sources

4b.1 **Are all the data elements available electronically?** (*elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims*)

☒ Yes ☐ No

4b.2 **If no, specify the near-term path to achieve electronic capture by most providers.**

N/A

Note: Measure stewards will be asked to specify the data elements for electronic health records at a later date

4b C☐ P☐ M☐ N☐

Comment [KP25]: 4b. The required data elements for the composite overall are available in electronic sources.

4d. Susceptibility to Inaccuracies, Errors, or Unintended Consequences

4d.1 **Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure and describe how these potential problems could be audited. If audited, provide results.**

Our measures are not susceptible to inaccuracies, errors, or unintended consequences; the component outcomes are well-specified in hospital administrative data.

4d C☐ P☐ M☐ N☐

Comment [KP26]: 4d. Susceptibility to inaccuracies, errors, or unintended consequences and the ability to audit the data items to detect such problems are identified.

4e. Data Collection Strategy/Implementation

4e.1 **Describe what you have learned/modified as a result of testing and/or operational use of the composite/component measures regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:**

Outcome component measures are derived from Medicare hospital claims, which are believed to be complete. All process component measures are reported as part of the Hospital Inpatient Quality Reporting Program in order for hospitals to receive the full annual Medicare payment update. Hospitals therefore have a strong financial incentive to provide process-of-care indicators. Continued availability of component measures for the AMI composite is therefore assured.

4.2 **Costs to implement the measure** (*costs of data collection, fees associated with proprietary measures*)**:**

The composite measure is calculated from process- and outcome-of-care indicators that are already publicly reported by hospitals. Hospitals and providers should not experience any additional costs or burden from the calculation of this measure.

4e.3 **Evidence for costs:** N/A

4e.4 **Business case documentation:** N/A

4e C☐ P☐ M☐ N☐

Comment [KP27]: 4e. Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, etc.) for obtaining all component measures can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use).

If the component measures are combined at the patient level, complete 4c.

4c. Exclusions

4c.1 **Do the specified exclusions require additional data sources beyond what is required for the numerator and denominator specifications?** ☐ No ☐ Yes ▶**If yes, provide justification**

4c H☐ M☐ L☐ N☐ NA☐

Comment [KP28]: 4c. Exclusions should not require additional data sources beyond what is required for scoring the measure (e.g., numerator and denominator) unless justified as supporting measure validity.

TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Feasibility*?

4

Steering Committee: Overall, to what extent was the criterion, *Feasibility*, **met?**

Rationale:

4 C☐ P☐ M☐

Rating: C=Completely; P=Partially; M=Minimally; N=Not at all; NA=Not applicable

	N☐
RECOMMENDATION	
Steering Committee: Do you recommend for endorsement? **Comments:**	Y☐ N☐ A☐

CONTACT INFORMATION

Co.1 **Measure Steward (Intellectual Property Owner)**
Organization: Centers for Medicare & Medicaid Services
Street Address: 7500 Security Boulevard, Mail Stop S3-02-01 **City:** Baltimore **State:** MD **ZIP:** 21244

Co.2 **Point of Contact: First Name:** Shaheen **Last Name:** Halim **Credentials (MD, MPH, etc.):** Ph.D., CPC-A
Email: Shaheen.Halim@cms.hhs.gov **Telephone:** (410) 786-0641 ext:

Co.3 **Measure Developer If different from Measure Steward**
Organization: Mathematica Policy Research
Street Address: 955 Massachusetts Avenue, Suite 801 **City:** Cambridge **State:** MA **ZIP:** 02139

Co.4 **Point of Contact: First Name:** Marian **Last Name:** Wrobel **Credentials (MD, MPH, etc.):** Ph.D.
Email: MWrobel@mathematica-mpr.com **Telephone:** 617-301-8971 ext:

Co.5 **Submitter**
Organization: ☐ Measure Steward ☐ Measure Developer
First Name: **Last Name:** **Credentials (MD, MPH, etc.):**
Email: **Telephone:** ext:

Co.6 **List any additional organizations that sponsored/participated in measure development:**

ADDITIONAL INFORMATION

Ad.1 **Workgroup/Expert Panel involved in measure development**
Provide a list of workgroup/panel member names and organizations. Describe the group's role in measure development.

On October 20, 2009, CMS convened an Advisory Panel on Medicare Education (APME) that included healthcare professionals involved with communication of quality information to consumers. CMS provided this panel with an overview of plans to include new composite measures on the Hospital Compare website, and solicited feedback from the group. In general, the group was supportive of CMS' plans to pursue composites and encouraged further development in this area.

APME Panel Members
Gwendolyn T. Bronson, SHINE/SHIP Counselor, Massachusetts SHINE Program
Yanira Cruz, Ph.D., President and Chief Executive Officer, National Hispanic Council on Aging
Nan-Kirsten Forté, Executive Vice President, Consumer Services, WebMD
Cathy C. Graeff, R.Ph., M.B.A., Partner, Sonora Advisory Group
Carmen R. Green, M.D., Professor, Anesthesiology and Associate Professor, Health, Management, and Policy, University of Michigan
Jessie C. Gruman, Ph.D., President, Center for Advancing Health
Cindy Hounsell, J.D., President, Women's Institute for a Secure Retirement
Gail Hunt, President and Chief Executive Officer, National Alliance for Caregiving
Deeanna Jang, Policy Director, Asian and Pacific Islander American Health Forum
Andrew Kramer, M.D., Professor of Medicine, Division of Health Care Policy and Research, University of Colorado, Denver
Sandy Markwood, Chief Executive Officer, National Association of Area Agencies on Aging
David W. Roberts, M.P.A., Vice President, Government Relations, Healthcare Information and Management System Society
Julie Bodën Schmidt, M.S., Associate Vice President, Training and Technical Assistance, National Association of Community Health Centers
Rebecca P. Snead, Chief Executive Officer and Executive Vice President, National Alliance of State Pharmacy

Associations and APME Chair

In 2006, CMS partnered with the Hospital Quality Alliance (HQA) in order to explore and assess strategies for improving the consumer friendliness of the Hospital Compare website. Staff representing the HQA principal organizations, which include the American Hospital Association, the Federation of American Hospitals, and the Association of American Medical Colleges, convened a working group charged with determining how to make Hospital Compare more consumer friendly over the short and long term. One of the key long-term recommendations from this group was to direct CMS/HQA to create condition- or procedure-specific composites related to current measures on Hospital Compare. Indeed, the group noted that such summary measures may help condense a large volume of information into a smaller, more manageable amount that is easier for decision-making.

Ad.2 If adapted, name of original measure: N/A
Ad.3 If adapted, original specifications ☐ **attachment or Ad.4 web page URL:**

Measure Developer/Steward Updates and Ongoing Maintenance
Ad.6 Year the measure was first released: N/A
Ad.7 Month and Year of most recent revision: N/A
Ad.8 What is the frequency for review/update of this measure? Annually
Ad.9 When is the next scheduled review/update for this measure? 2012

Ad.10 Copyright statement/disclaimers:

Ad.11 Additional Information ☐ **attachment or web page URL:**

I have checked that the submission is complete and all the information needed to evaluate the measure is provided in the form; any blank fields indicate that no information is provided. ☒

Date of Submission (*MM/DD/YY*)**:** 11/3/10