THE NATIONAL QUALITY FORUM

COMPOSITE MEASURE SUBMISSION FORM Version 4.1 January 2010

This form will be used by stewards to submit <u>composite</u> measures and by reviewers to evaluate the measures.

Measure Stewards: Check with NQF staff before using this form. Complete all <u>non-shaded</u> areas of the form. All requested information should be entered directly into this form. The information requested is directly related to NQF's <u>composite measure evaluation criteria</u> 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.

<u>Note</u>: 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 #: 962 NQF Project:				
De.1 Title of Measure: Composite Measure of Hospital Quality for Heart Failure (HF)				
De.2 Brief description of measure (<i>including type of score</i> , <i>measure focus</i> , <i>target population</i> , <i>time</i> , <i>e.g.</i> , <i>Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year</i>): A composite measure of in-hospital process- and outcome-of-care for Heart Failure (HF) 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				

				NQF F	Review #:
De.5 IOM Quality Domain 🔀 effectiveness 🔀 timeliness	efficiency	equity	patient-centered	safety	
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 (<u>measure steward agreement</u>) is signed. <i>Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.</i>	
A.1 Do you attest that the measure steward holds intellectual property rights to the measure <u>and</u> the right to use any aspects of the measure owned by another entity (e.g., component measures, risk model, code set)? Xes	
A.2 Measure Steward Agreement Signed and Submitted OR Sovernment entity-public domain (If measure steward agreement not signed for non-government entities, do not submit)	A Y□ N□
A.3 Please check if either of the following apply: Proprietary Measure Proprietary Complex Measure w/fees	
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 <u>both</u> public reporting <u>and</u> quality improvement. C.1 Purpose: Public reporting Internal quality improvement C.2 Accountability Accreditation Payment incentive Other, describe: (If not intended for <u>both</u> public reporting <u>and</u> 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: X Fully developed and tested (If composite measure not tested, do not submit)	D Y□
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.	N
► Is all requested information entered into this form? Yes (If no, do not submit)	
endorsed or submitted for consideration for NQF endorsement (<i>check one</i>) All component measures are <u>NQF-endorsed</u> measures <u>Some or all</u> component measures are <u>not NQF-endorsed</u> 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):	

TAP/Workgroup Reviewer Name:

Steering Committee Reviewer Name:	
1. IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in or overall poor performance. <i>Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria.</i> (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 various measures adopted for the program 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 Heart Failure (HF) 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 discussion of the construction of formative composite measures appears in Appendix B.	
Specifically, this measure summarizes clinical process- and outcome-of-care indicators associated with the treatment of HF 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 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 may 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 assess: "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 HF 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 four individual process-of-care indicators and two outcome-of-care indicators meant to capture the quality of hospital care provided to patients with HF. NQF has endorsed all six indicators. The proposed composite combines these in the form of process- and outcome-of-care domain scores. CMS realizes that some HF indicators that appear <i>on Hospital Compare</i> and are included in the composite measure may later lose their endorsed status. Should that occur, we will reconfigure the composite and resubmit to NQF for endorsement at the next available opportunity.	1d C P M N

However, CMS wishes the composite to include all HF indicators that are endorsed at the time of its submission.

CMS developed the composite measure to achieve the following goals for reporting hospital quality measures composite methodology:

- 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 focuses on diseases important to the Medicare population: Acute Myocardial Infarction (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 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, we present the mapping between CMS' policy goals and methodological decisions in tabular form.

Policy Goals	Methodological Decisions
Summarize measures on Hospital Compare in a single, useful, condition-specific composite	 Include the same set of process and outcome measures as Hospital Compare
Produce differences in composite values that are clinically and statistically meaningful and reflect true differences in underlying quality	 Enter component indicators as values, not ranks, so that slight differences in measured performance do not potentially lead to large differences in the composite value for topped-off measures For process indicators, adjust component indicators for reliability so that random variation does not drive small hospitals to extremes
Results available for a large number of hospitals	• Process indicators are available when the number of eligible discharges is five or more; outcome indicators are available when the number of eligible discharges is 25 or more
Focus more on measures relevant to more patients	 Construct process and outcome domains using weights based on national denominators
Method is scientifically acceptable and acceptable to stakeholders	 Adopt an approach that is similar to that used for AHRQ quality indicators (QIs) Note: AHRQ QIs are NQF-endorsed and widely reported
Method accommodates changes in the set of measures on Hospital Compare	
Method can be used for multiple conditions	 Method is based on general principles, not on the specific statistical performance of a group of measures Process and outcome domains are statistically standardized
Relative weighting of process and outcome domains does change when measures are added to or deleted from one domain	before they are added together
Method is relatively simple Hospitals can duplicate results	 Use equal weighting to combine process and outcome domains Reliability weights are a function of a hospital's number of cases and national parameters

Table 1d.2.1. CMS Policy Goals for Composite Measures and Associated Methodological Decisions

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, the HF composite is a formative summary of all HF indicators reported on Hospital Compare. Measures were adopted for the Hospital Inpatient Quality Reporting Program through a consensus process that deemed them to be indicators of well-coordinated high-quality care for HF. The measures that make up the composite include both process- and outcome-of-care indicators.

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 the HF composite measure, the process-of-care indicators

1	е
С	
Ρ	
Μ	
Ν	

evaluate whether a patient received:

- Discharge instructions [HF1]
- Evaluation of Left Ventricular Systolic (LVS) Function [HF2]
- ACE Inhibitor or ARB for Left Ventricular Systolic Dysfunction (LVSD) [HF3]
- Smoking Cessation advice/counseling [HF4]

These NQF-endorsed process-of-care indicators represent established best practices for HF care^{1,2} and were adopted by CMS 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 HF patients³⁻⁵, one of the two outcome-of-care indicators also included in the proposed HF composite measure. The two HF outcome-of-care component measures are: 1) 30-day risk-standardized mortality and 2) 30-day risk-standardized 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 HF patients at a given hospital experience, and therefore may be critical to understanding the quality of care received.ⁱ

The combination of these component indicators ultimately serves to deliver a single, useful, condition-specific summary of HF care for consumer use.

Citations

- 1. Hunt SA. ACC/AHA 2005 guideline update for the diagnosis and management of chronic heart failure in the adult: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Writing Committee to Update the 2001 Guidelines for the Evaluation and Management of Heart Failure). J American College of Cardiology 2005; 46(6):e1-82.
- 2. Heart Failure Society of America. HFSA 2006 Comprehensive Heart Failure Practice Guideline. J Card Fail. 2006 Feb; 12(1):e1-2.
- 3. 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.
- 4. Garg R, Yusuf S. Overview of randomized trials of angiotensin-converting enzyme inhibitors on mortality and morbidity in patients with heart failure. Collaborative Group on ACE Inhibitor Trials. JAMA. 1995 May 10; 273(18):1450-6. Erratum in: JAMA 1995 Aug 9; 274(6):462.
- 5. Heart Failure Society of America. HFSA 2010 Comprehensive Heart Failure Practice Guideline. J Card Fail. 2010 Jun; 16(6):e1-e2.

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

If the component measures are <u>combined at the aggregate level</u>, 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 numl	bers	frequently	performed procedure	re 🛛 🗌 leading cause of morbidity/mortality	/ 🗌
high resource use	severity	y of illness	patient/societal	al consequences of poor quality	
other, describe: 1	a.2				

1a.3 Summary of Evidence of High Impact:

1a.4 Citations for Evidence of High Impact:

1a H[

MΓ

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.

1b. Opportunity for Improvement 1b.1 Briefly explain benefits (improvements in quality) envisioned by use of this measure:	
1b.2 Summary of data demonstrating performance gap (<i>variation or overall poor performance across providers</i>):	
1b.3 Citations for data on performance gap:	1b
1b.4 Summary of Data on disparities by population group:	
1b.5 Citations for data on Disparities:	
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 (<i>also provide narrative description of the rating and by whom</i>) 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 (<i>including guideline number and/or page number</i>)	
1c.10 Clinical Practice Guideline Citation: 1c.11 National Guideline Clearinghouse or other URL:	
1c.12 Rating of strength of recommendation (<i>also provide narrative description of the rating and by whom</i>)	10
1c.13 Method for r ating strength of recommendation (<i>If different from <u>USPSTF system</u></i> , also describe rating and how it relates to USPSTF):	
1c.14 Rationale for using this guideline over others:	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Importance to Measure and Report?</i>	1
Steering Committee: Was the threshold criterion, <i>Importance to Measure and Report</i> , 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:	2a- specs C P

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

M N

http://www.hospitalcompare.hhs.gov/ S.2 If yes, provide web page URL: http://www.hospitalcompare.hhs.gov/ 2a. Precisely Specified **2a.0.1** Components of the Composite (List the components, i.e., domains/sub-composites, individual measures. If component measures are NOF-endorsed, include NQF measure number; if not NQF-endorsed, provide date of submission to NQF) HOSPITAL PROCESS-OF-CARE INDICATORS NQF #0136 Percent of HF Patients that Received Discharge Instructions Endorsed May 9, 2007 Percent of HF Patients with Evaluation of LVS Function Endorsed May 9, 2007 NOF #0135 NQF #0162 Percent of HF Patients Given ACE Inhibitor or ARB for LVSD Endorsed May 9, 2007 NQF #0027 Percent of HF Patients Given Smoking Cessation Advice/Counseling Endorsed May 1, 2006 HOSPITAL OUTCOME-OF-CARE INDICATORS NQF #0229 HF 30-day Risk-Standardized Mortality Endorsed May 9, 2007 NQF #0330 HF 30-day Risk-Standardized Readmission Endorsed May 15, 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: For the process-of-care domain, the numerator is equal to the weighted sum of four terms. Each term is equal to the ratio of the hospital's raw performance rate to the national performance rate for the indicator. The weight is equal to the total number of observations, that is, the number of patients 'at risk' for the indicator. For the outcome-of-care domain, the numerator is equal to the weighted sum of two terms. Each term is equal to the ratio of the hospital's risk-standardized performance rate to the national performance rate for the indicator. The weight is equal to the total number of eligible discharges for the indicator. 2a.2 Numerator Time Window: July 2006 - June 2009 **2a.3 Numerator Details:** Successes in the following heart failure process-of-care and outcome-of-care indicators: HOSPITAL PROCESS-OF-CARE INDICATORS Percent of HF Patients that Received Discharge Instructions (NQF #0136) 1. Percent of HF Patients with Evaluation of LVS Function (NQF #0135) 2. Percent of HF Patients Given ACE Inhibitor or ARB for LVSD (NQF #0162) 3. Percent of HF Patients Given Smoking Cessation Advice/Counseling (NQF #0027) 4. HOSPITAL OUTCOME-OF-CARE INDICATORS HF 30-day Risk-Standardized Mortality (NQF #0229) 1. 2. HF 30-day Risk-Standardized Readmission (NQF #0330) 2a.4 Composite Denominator Statement: For the process-of-care domain, the denominator is equal to the total number of observations for all HF process indicators. It is thus equal to the number of patients 'at risk for the four process indicators. For the outcome-of-care domain, the denominator is equal to the total number of observations for all HF outcome indicators. It is thus equal to the number of eligible discharges for the two outcome indicators. 2a.5 Target Population Gender 🔀 Female X 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 heart failure quality reports. Counts of outcome-of-care opportunities are based on claims data.

2a.9 Composite Denominator Exclusions:

The following two criteria were applied as exclusion restrictions:

- 1. Hospitals with less than five eligible patient cases for the process-of-care indicators and less than 25 eligible discharges for the outcome-of-care indicators.
- 2. Hospitals that were missing rates for one or more process-of-care and/or outcome-of-care indicators.

2a.10 Denominator Exclusion Details: See above (2a.9)

2a.11 Stratification Details/Variables (*All information required to stratify the measure including the stratification variables, all codes, logic, and definitions***)**: The composite measure was not stratified.

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 as the simple average of process and outcome domain scores. The outcome domain score was computed as the denominator-weighted sum of the ratio of actual to expected values of the two outcome indicators. The process domain score was computed as the ratio of actual to expected values of the four process indicators. All indicators are publically reported by the CMS on *Hospital Compare* and are NQF endorsed. The method of scoring is described in detail below. Additional documentation is available in Section 2 of the attached appendix (Appendix A).

CMS began publically reporting 30-day risk-standardized mortality and readmission rates, used in construction of the outcome domains score, in June 2007 and in July 2009, respectively. In computing the indicators, Yale researchers employed a method known as 'shrinkage' or 'Bayesian smoothing' to increase the overall accuracy of the indicators. The method is well-known and widely accepted in the statistical literature (Morris 1983; Carlin and Louis 2000). In order to bring the process-of-care indicators into conformity with outcome indicators in constructing the composite, reliability weights to each individual process-of-care indicator. Each indicator is thus computed as a weighted average of the hospital's own value for the indicator and the national mean for that indicator. Each indicator was then standardized by dividing by the national mean of the indicator. Outcome-indicators were also was standardized by dividing by the national mean of the indicator.

In order to remain consistent with the approach used for AHRQ measures, CMS used denominator weighting in constructing the process- and outcome-of-care domains. Denominator weighting places greater weight on indicators that apply to higher numbers of patients nationally, 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-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.

Lastly, the overall composite score was calculated as a simple average of the two domain scores. In Table 2a.42.1, we provide a summary of the composite measure. Since the process- and outcome-of-care indicators are standardized by the national rate of each of the indicators, hospitals with a composite score of >1 have a performance score that is greater than the national rate and hospitals with a composite score of <1 have a performance score that is less than the national rate. However, it should be noted that the differences in performance from the national rate should be interpreted with caution since it may not be statistically significant.

Therefore, our method of discrimination of performance is described in greater detail in Section 2a.22.					
Table 2a.42.1: Summary of Composite and Composite Domains					
Domain	Description	Interpretation			
Process-of-Care	Denominator weighted average of standardized (by the national mean) probabilities that patients with HF will	Hospitals with a process-of-care domain score >1 have a score that is better than average.			
	receive the appropriate care.	Hospitals with a process-of-care domain score <1 have a score that is worse than average.			
		Hospitals with a process-of-care domain score =1 have a score that is equal to the average.			
Outcome-of-Care	Denominator weighted average of standardized (by the national mean) probabilities of survival and of avoidance of readmission after 30 days of admission to a hospital with HF.	Hospitals with an outcome-of-care domain score >1 have a score that is better than average. Hospitals with an outcome-of-care domain			
		score <1 have a score that is worse than average.			
		Hospitals with an outcome-of-care domain score =1 have a score that is equal to the average.			
Overall Composite	Simple average of the process- and outcome-of-care domain scores.	Hospitals with a composite score >1 have a score that is better than average.			
		Hospitals with a composite score <1 have a score that is worse than average.			
		Hospitals with a composite score =1 have a score that is equal to the average.			

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

Composite scores for a hospital were calculated if:

- 1. The hospitals reported rates for all four process and all two outcome-of-care indicators
- 2. Each process-of-care indicator had at least five cases and each outcome-of-care indicator had at least 25 cases.

Composite scores were not estimated for hospitals that did not satisfy the above two criteria. Table 2a.44.1 summarizes the time at which the data was released on *Hospital Compare* and the collection period of the quality indicators. In addition, Figure 2a.44.1 shows how the final sample of hospitals was derived.

Table 2a.44.1: Data Release and Collection Period

Data Release	Indicators Used	Time Period
March 2008	Process-of-care	July 2006-June 2007
March 2009	Process-of-care	July 2007-June 2008
March 2010	Process-of-care	July 2008-June 2009
June 2010	Outcome-of-care	July 2006-June 2009



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

Consistent with the approach used for the AHRQ measures, CMS used denominator weighting in constructing the process- and outcome-of-care domains. Denominator weighting places relatively more weight on measures that apply to relatively more patients nationally, 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 NOF 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. Technical documentation on the scoring approach is provided in Section 2.1 of Appendix A, attached)

2a.21 Calculation Algorithm (Describe the calculation of the measure as a flowchart or series of steps):

Key Steps Process-of-Care Domain Outcome-of-Care Domain Overall Composite Step 1a Exclude hospitals Exclude hospitals if there are less than five cases for any of that do not meet the minimum case size requirement Exclude hospitals if there are less than 125 cases for any of the two outcome-of-care indicators. Exclude hospitals if there are less than 25 cases for any of the two outcome-of-care indicators. Step 1b Exclude hospitals more indicators Exclude hospitals missing one or more process-of-care indicators. Exclude hospitals missing one or more outcome-of-care N/A Step 2 Weight the indicators is set to a weighted average of the hospital's own rate and the national rate. The value of each process-of- care indicator is set to a weighted average of HF patients with evaluation of LVS function" at Heartcare Regional Hospital's 80% and the national rate for this indicator is 77%. Also, suppose that the hospital's weight is 0.8. Then the hospital's ones the performance rate for the sopital's one ach the national rate is: 0.8(80%) + (1 - 0.8)(77%) = 79.4% N/A N/A Step 3 The value of each (reliability weight ig 0.2.8. Then the hospital's reliability-weight adjusted rates is: 0.8(80%) + (1 - 0.8)(77%) = 79.4% The value of each outcome-of- care indicator is divided by the national rate.	Table 2a.21.1: Steps to Construct the Composite Score					
Step 1a Exclude hospitals that do not meet the minimum case size requirement Exclude hospitals if there are less than five cases for any of the four process-of-care indicators. Exclude hospitals if there are less than 25 cases for any of the two outcome-of-care indicators. Step 1b Exclude hospitals Exclude hospitals Exclude hospitals more indicators Exclude hospitals missing one or more process-of-care indicators. Exclude hospitals missing one or more outcome-of-care indicators. Step 2 Weight the indicators by a reliability weight The value of each process-of- care indicators is set to a weighted average of the hospital's own rate and the national rate. Exclude hospitals missing one or more outcome-of-care indicators. Example Suppose the performance rate for the "percentage of HF patients with evaluation of LVS function" at Heartcare Regional Hospital's s08 and the national rate for this indicators is 37%. Also, suppose that the hospital's adjusted rates is: 0.8(80%) + (1 - 0.8)(77%) = 79.4% N/A N/A Step 3 standardize the indicators by dividing by the national rate. The value of each (reliability weight adjusted) proces-of- care indicator is divided by the national rate. The value of each outcome-of- care indicator is divided by the national rate.	Key Steps	Process-of-Care Domain	Outcome-of-Care Domain	Overall Composite		
Step 1b Exclude hospitals missing one or more indicatorsExclude hospitals missing one or more process-of-care indicators.Exclude hospitals missing one or more outcome-of-care indicators.Step 2 Weight the indicators by a reliability weightThe value of each process-of- care indicators is set to a weighted average of the hospital's own rate and the national rate.Image: Comparison of the patients with evaluation of LVS function" at Heartcare Regional Hospital's 0.8. Then the hospital's reliability-weight adjusted rates is: 0.8(80%) + (1 - 0.8)(77%)) = 79.4%N/AN/AStep 3 Standardize the indicators by dividing by the national mean ofThe value of each (reliability weight adjusted) process-of- care indicator is divided by the national rate.The value of each outcome-of- care indicator is divided by the national rate.	Step 1a Exclude hospitals that do not meet the minimum case size requirement	Exclude hospitals if there are less than five cases for any of the four process-of-care indicators.	Exclude hospitals if there are less than 25 cases for any of the two outcome-of-care indicators.	N/A		
Step 2 Weight the indicators by a reliability weight The value of each process-of- care indicators is set to a weighted average of the hospital's own rate and the national rate. Example Suppose the performance rate for the "percentage of HF patients with evaluation of LVS function" at Heartcare Regional Hospital is 80% and the national rate for this indicator is 77%. Also, suppose that the hospital's weight is 0.8. Then the hospital's reliability-weight adjusted rates is: 0.8(80%) + (1 - 0.8)(77%) = 79.4% N/A N/A Step 3 Standardize the indicators by dividing by the national mean of The value of each (reliability weight adjusted) process-of- care indicator is divided by the national rate. The value of each outcome-of- care indicator is divided by the national rate.	<u>Step 1b</u> Exclude hospitals missing one or more indicators	Exclude hospitals missing one or more process-of-care indicators.	Exclude hospitals missing one or more outcome-of-care indicators.			
Step 3The value of each (reliability weight adjusted) process-of- care indicator is divided by the national rate.The value of each outcome-of- care indicator is divided by the national rate.Standardize the indicators by dividing by the national mean ofThe value of each outcome-of- care indicator is divided by the national rate.Example	Step 2 Weight the indicators by a reliability weight	The value of each process-of- care indicators is set to a weighted average of the hospital's own rate and the national rate. Example Suppose the performance rate for the "percentage of HF patients with evaluation of LVS function" at Heartcare Regional Hospital is 80% and the national rate for this indicator is 77%. Also, suppose that the hospital's weight is 0.8. Then the hospital's reliability-weight adjusted rates is: 0.8(80%) + (1 - 0.8)(77%) = 79.4%	N/A	N/A		
each indicatorExample Given the previous example in Step 2, if Heartcare Regional Hospital's reliability-weight adjusted rates is 79.4% and the national reliability-rate adjusted rate is 81%, then the standardized indicator is: $\frac{79.4}{81.0} = 0.98$ If the 30-day risk-adjusted survival rate at Heartcare Regional Hospital is 91% and the national survival rate is 88.8%, then the standardized indicator is: $\frac{91.0}{88.8} = 1.02$ N/A	Step 3 Standardize the indicators by dividing by the national mean of each indicator	The value of each (reliability weight adjusted) process-of- care indicator is divided by the national rate. Example Given the previous example in Step 2, if Heartcare Regional Hospital's reliability-weight adjusted rates is 79.4% and the national reliability-rate adjusted rate is 81%, then the standardized indicator is: $\frac{79.4}{81.0} = 0.98$	The value of each outcome-of- care indicator is divided by the national rate. Example If the 30-day risk-adjusted survival rate at Heartcare Regional Hospital is 91% and the national survival rate is 88.8%, then the standardized indicator is: $\frac{91.0}{88.8} = 1.02$	N/A		

Table 2a.21.1: Steps to Construct the Composite Score (cont.)						
Key Steps	Process-of-Care Domain	Outcome-of-Care Domain	Overall Composite			
<u>Step 4</u> Combine the indicators using a denominator weighted average	Take a denominator- weighted average of the standardized process-of-care indicators.Example Suppose the standardized rates and the national number of cases for the four process-of-care for Heartcare Hospital respectively are*: HF1: 1.10 (N=4000) HF2: 0.98 (N=5000) HF3: 1.32 (N=3500) HF4: 0.95 (N=4000)Then the process-of-care domain score is: $\frac{4000}{16500}(1.10) + \frac{5000}{16500}(0.98)$ 	Take a denominator- weighted average of the standardized outcome-of- care indicators.Example Suppose the standardized rates and the national number of cases for the two outcome-of-care for Heartcare Hospital respectively are**: Survival: 1.02 (N=4500) Readmission: 0.95 (N=4500)Then the outcome-of-care domain score is: $\frac{4500}{9000}(1.02) + \frac{4500}{9000}(0.95)$ = 0.99	N/A			
Step 5 Combine the process- and outcome-of-care domains to create a composite score	N/A	N/A	Take a simple average of the process- and outcome-of- care domain scores Example Given the standardized rates for the process- and outcome-of-care domains, the composite score is: $\frac{1}{2}(1.06) + \frac{1}{2}(0.99) = 1.03$			

Notes:

* HF1: Percent of HF Patients that Received Discharge Instructions; HF2: Percent of HF Patients with Evaluation of LVS Function; HF3: Percent of HF Patients Given ACE Inhibitor or ARB for LVSD; HF4: Percent of HF Patients Given Smoking Cessation Advice/Counseling.

** Survival: 30-day risk-adjusted survival rate; Readmission: 30-day risk-adjusted lack of readmission.

2a.22 Describe the method for discriminating performance (*e.g.*, *significance testing*):

To examine meaningful differences in composite measures among hospitals, we compared hospitals' confidence interval estimates with the overall mean and assigned hospitals into one of three performance categories: "better-than-expected' hospitals, if the interval estimate is entirely above the mean; 'no-different-thanexpected' hospitals, if the interval estimate includes the mean; and 'worse-than-expected' 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.

We derived the standard error for each hospital and estimated 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. Figure 2a.22.1 shows how the hospitals are categorized into one of three performance categories. Complete information on the technical methodology for discriminating performance is contained in Appendix A, Section 2.3.



Figure 2a.22.1: Hospital Categorization

NQF Review #:

2a.26 Data source/data collection instrument attached http://www.hospitalcompare.hhs.gov/	OR 2a.27 at web page URL:
2a.29 Data dictionary/code table attached OR 2a.30 http://www.hospitalcompare.hhs.gov/	at web page URL:
2a.32 Level of Measurement/Analysis (Check the level for	or which the measure is specified and tested)
Clinicians: Individual Group Other Facility/Agency (e.g., hospital, nursing home) Health plan Integrated delivery system Multi-site/corporate chain Population: National Regional/network	 Prescription drug plan Program: Disease management QIO Other Measured at all levels
State Counties/Cities	Other (<i>Please describe</i>):
2a.26 Care Settings (<i>Check the settings for which the me</i> Ambulatory Care: Amb Surgery Center Office	easure is specified and tested; check all that apply) 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 (<i>Please describe</i>):
2a.38 Clinical Services (Healthcare services being measure	red; 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 national	evel and include outcomes, complete the following
2a.12 Risk Adjustment Type: No risk adjustment nece adjustment paired data at patient level risk-adju risk adjustment method widely or commercially availabl Other (specify) 2a.13	ssary analysis by subgroup case-mix ustment devised specifically for this measure/condition e
2a.14 Risk Adjustment Methodology/Variables (<i>List risk statistical models, or other aspects of model or method</i>)	adjustment variables and describe conceptual models,
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 HF that are now reported under the Hospital Inpatient Quality Reporting Program. Our analysis aims to document the strength of associations among them.

The analysis reported here relies on data that are publicly reported on Hospital Compare. We merged process-ofcare indicators and outcome-of-care indicators for HF collected between July 2006 and June 2009. We estimated composite measures for 3,586 hospitals (out of a potential 4,240 hospitals) for which:

- 1. The hospitals reported rates for all four process and all two outcome-of-care indicators
- 2. Each process-of-care indicator had at least five cases and each outcome-of-care indicator had at least 25 cases.

Background on Indicators Reported on Hospital Compare:

The indicators used in the construction of composites were drawn from *Hospital Compare*. The process-of-care indicators were drawn from Medicare hospital administrative claims data and medical record documents with discharge dates between July 2006 and June 2009. The hospital outcome-of-care indicators for 30-day risk-adjusted mortality and readmission for HF were based on Medicare claims for hospital stays with discharge dates between July 2006.

2i.2 Analytic Method:

We carried out two analyses to explore the structure of the HF 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 Tables 2i.3.1 and 2i.3.2

2i.3 Results:

Although the HF composite was not intended as a reflective measure, psychometric properties do indicate a single underlying quality construct.

Table 2i.3.1 shows correlations across the process and outcome indicators. The correlations across the process-ofcare indicators are significant and positive, and all are greater than 0.4, which indicates moderate correlation. Correlations between the process and outcome indicators are positive, albeit are weak, with values below 0.10. There is a weak negative correlation between mortality and readmission, which may reflect competing risks. That is, higher rates of mortality reduce the opportunity for readmission. Cronbach's alpha was estimated as 0.73, surpassing the commonly desired value of 0.70, suggesting that indicators are internally consistent.

The factor analysis of component measures produced a single factor with an eigenvalue greater than one. The eigenvalue for the first factor was almost 10 times that of the second factor, strongly suggesting that the component indicators represent one underlying construct.



Table 2i.3.1. Correlation of Variables in HF Composite Measure

	HF 1	HF 2	HF 3	HF 4	Mort	Read
HF 1	1.00					
HF 2	0.47	1.00				
HF 3	0.40	0.51	1.00			
HF 4	0.59	0.51	0.53	1.00		
Mort*	0.07	0.05	0.10	0.18	1.00	
Read*	0.09	0.03	0.03	0.07	-0.13	1.00
Cronbach Alpha	0.73					

Notes:

* Mort: Survival rate, where Mort=100-(30-day risk-standardized mortality rate); Read: absence of readmission, where Read=100-(30-day risk-standardized readmission rate).

Table 2i.3.2. Factor Analysis Results

	Fa	actor Loading	js	
	Factor 1	Factor 2	Factor 3	Uniqueness
HF 1	0.63	0.05	0.11	0.59
HF 2	0.73	0.01	-0.08	0.47
HF 3	0.75	-0.03	-0.09	0.43
HF 4	0.78	-0.02	0.06	0.38
Mort*	0.10	-0.30	0.06	0.90
Read*	0.08	0.30	0.04	0.90
Eigenvalues	2.12	0.19	0.04	
Proportion	1.12	0.10	0.02	
Ν	3,586			

Notes:

* Mort: Survival rate, where Mort=100-(30-day risk-standardized mortality rate); Read: absence of readmission, where Read=100-(30-day risk-standardized readmission rate).

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

2j.1 Data/sample:

The analysis of the component indicators' contribution to variability of the composite relies on data that are publicly reported on Hospital Compare. We merged process-of-care indicators and outcome-of-care indicators for HF collected between July 2006 and June 2009. We estimated composite measures for 3,586 hospitals (out of potential 4,240 hospitals) for which:

- 1. The hospitals reported rates for all four process and all two outcome-of-care indicators
- 2. Each process-of-care indicator had at least five cases and each outcome-of-care indicator had at least 25 cases.

<u>Background on Indicators Reported on *Hospital Compare*: The indicators used in the construction of composites were drawn from *Hospital Compare*. The process-of-care</u>

2j

РΓ

M

N[

indicators were drawn from Medicare hospital administrative claims data and medical record documents with discharge dates between July 2006 and June 2009. The hospital outcome-of-care indicators for 30-day risk-adjusted mortality and readmission for HF were based on Medicare claims for hospital stays with discharge dates between July 2006 and June 2009.

2j.2 Analytic Method:

In order to assess the contribution of each indicator to variability in the HF composite, we compare the percent change in (1) the variance and (2) the inter-quartile range (IQR) of the composite and of the process and outcome domain scores when a process or outcome indicator is removed. Results appear in Table 2j.3.1.

2j.3 Results:

In Table 2j.3.1, positive values indicate that addition of the component indicator tends to *reduce* the variance or IQR. Only one indicator, HF2 (Percent of HF Patients with Evaluation of LVS Function), exhibits a positive effect on the composite variance. Because the outcome domain contains only two component indicators, readmission and mortality both have strong negative effects on the variance of the domain score. The strong variance-reducing effect of mortality appears to be the result of its tight distribution.

	Overall C	omposite	Process	Domain	Outcome	e Domain
		Change in		Change in		Change in
		Inter-		Inter-		Inter-
	Change in	quartile	Change in	quartile	Change in	quartile
	Variance	Range	Variance	Range	Variance	Range
Remove:	(%)	(%)	(%)	(%)	(%)	(%)
HF 1	21.10	7.93	21.78	8.29	-	-
HF 2	-32.83	-33.84	-34.22	-36.41	-	-
HF 3	4.65	4.86	4.77	5.16	-	-
HF 4	42.62	36.55	44.66	37.92	-	-
Mortality	2.54	2.04	-	-	194.17	72.82
Readmission	0.09	1.28	-	-	25.83	17.24

Table 2j.3.1. Change in Inter-quartile Range and Variance of the Composite, Process and Outcome Domains with the Removal of Indicators

2k. Analysis to support differential weighting of component scores

2k.1 Data/sample:

In constructing the composite, individual component indicators are weighted, in each instance, by the national 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, though this effect might be muted if individual indicators have similar distributions.

Testing to support differential weighting of composite uses data that are publicly reported on Hospital Compare. We merged process indicators and outcome indicators for HF collected between July 2006 and June 2009. We estimated composite measures for 3,586 hospitals (out of potential 4,240 hospitals) for which:

- 1. The hospitals reported rates for all four process and all two outcome-of-care indicators
- 2. Each process-of-care indicator had at least five cases and each outcome-of-care indicator had at least 25 cases.

Background on Indicators Reported on Hospital Compare:

The indicators used in the construction of composites were drawn from Hospital Compare. The process-of-care indicators were drawn from Medicare hospital administrative claims data and medical record documents with discharge dates between July 2006 and June 2009. The hospital outcome-of-care indicators for 30-day risk-adjusted mortality and readmission for HF were based on Medicare claims for hospital stays with discharge dates between July 2006.



2k.2 Analytic Method:

We compared the distribution of the HF composite measure with equal and differential weighting.

2k.3 Results:

Figure 2k.3.1 displays the distribution of the HF composite measure with equal and differential weighting. As the figure shows, denominator weighting slightly increases the percentage of hospitals with higher composite scores. A table of the distribution of composite scores is also provided in the appendix (Table 2k.3.1)





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, not 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 large numbers 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 HF 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 HF 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, we tested an alternative scoring approach that differed on two levels (Alternate Method). First, we estimated composite scores for hospitals that were missing less than half of

the process- and outcome-of-care indicators. That is, if a hospital had two or more process and one or more outcome indicator, a composite score was estimated. We imputed missing values with the national mean. Second, we used an alternative standardization approach by subtracting the national mean and dividing by the standard deviation, before taking the simple average of the two domain scores. Because this could result in negative composite values for some hospitals, the score was then rescaled to a range between zero and one hundred. It should be noted that this approach was a method we used when we initially presented our composite measure to the NQF in February, 2011.

In Figure 2k.5.1, we present distributions of the two alternative scoring methods. The figures show that the second approach (Alternate Method) leads to composite scores with a tight distribution as a result of the standardization approach; therefore, our proposed approach should provide users with a distribution that is easier for consumers to view. Furthermore, our reevaluated compositing approach reduces potential misinterpretations by consumers that the composite score is an actual rate between zero and 100 percent. A table of the distribution of composite scores is also provided in the appendix (Table 2k.5.1)



Figure 2k.5.1: Comparision of Compositing Approaches

Furthermore, we considered, but rejected, alternative weighting schemes that would reduce the weight assigned to indicators that were strongly left-skewed (often referred to as "topped off"). This can be done, for example, by constructing weights that depend on the difference between the national mean for an indicator and the highest possible score. First, we are disinclined to make judgments about the relative importance of endorsed indicators. It does not appear reasonable to argue that an element of care becomes "less important" in a composite because many hospitals report providing it. Second, at a purely practical level, the distributions of the four HF process indicators do not sharply differ from one another, so weighting in this fashion would produce a result resembling equal weighting. Finally, and perhaps most importantly, such an approach to weighting would make a hospital's score dependent on the behavior of other hospitals. For example, a hospital that performed well on indicator A and poorly on indicator B would receive a higher score if other hospitals performed poorly on A and well on B than it would if other hospitals performed well on A and poorly on B. This is not, in our view, a desirable property for a composite to have.

21. Analysis of missing component scores

2I.1 Data/sample:

Construction of the composite scores relies on data that are publicly reported on Hospital Compare. We merged process-of-care indicators and outcome-of-care indicators for HF collected between July 2006 and June 2009. We estimated composite measures for 3,586 hospitals (out of potential 4,240 hospitals) for which:

- 1. The hospitals reported rates for all four process and all two outcome-of-care indicators
- 2. Each process-of-care indicator had at least five cases and each outcome-of-care indicator had at least 25 cases.

21

СΓ

Ρſ

M N Of the 4,240 hospitals 654 did not receive a composite score for one or more of the following reasons:

- 1. The hospital was missing a rate for one or more of the process- and/or outcome-of-care indicators (1.2%)
- 2. The hospital reported a case size of zero for one or more of the process-of-care indicators; therefore a hospital specific rate was not reported (3.8%)
- 3. The hospital reported a case size of greater than zero, but less than five cases for one or more process-of-care indicator (1.2%)
- 4. The hospital reported a case size of less than 25 cases for one or more outcome-of-care indicator (0.8%)

Background on Indicators Reported on Hospital Compare:

The indicators used in the construction of composites were drawn from *Hospital Compare*. The process-of-care indicators were drawn from Medicare hospital administrative claims data and medical record documents with discharge dates between July 2006 and June 2009. The hospital outcome-of-care indicators for 30-day risk-adjusted mortality and readmission for HF were based on Medicare claims for hospital stays with discharge dates between July 2006.

2I.2 Analytic Method:

We examined whether there were differences in the distribution of the process- and outcome-of care rates for all hospitals compared to those hospitals for which there were no missing process- and outcome-of-care indicators so that composites were estimated for these hospitals.

2I.3 Results:

Figures 2I.3.1 and 2I.3.2 show that there is very little difference in the distribution of each of the components indicators between those hospitals that had a composite score calculated (i.e., those with no missing process- or outcome-of-care indicators and for the full sample of hospitals. Specific distributions for each of the indicators are available in Table 2I.3.1 in the appendix.



Figure 21.3.1: Comparison of Distribution for Process-of-Care Indicators

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



ⁱⁱ 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):

1. 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 percent agreement for continuous variable elements and chancecorrected agreement using Cohen's kappa for binary data elements.

2. 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% 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):

1. 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 below summarizes the reliability statistics for the HF 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 HF-3 (ACEI for LVSD), substantial agreement (kappa ranging from 0.61 - 0.80) for HF-2 (LVSD evaluation) and HF-4 (smoking cessation), and moderate agreement (kappa ranging from 0.41 - 0.60) for HF-1 (discharge instructions).

HF Component Measure	N	Agreement (%)	Карра
HF-1*			
Discharge instructions to address activity	180	86.1	0.65
Discharge instructions to address diet	180	90.0	0.73
Discharge instructions address follow-up	180	87.8	0.47
Discharge instructions address medications	180	90.6	0.53
Discharge instructions address symptoms	180	86.1	0.71
Discharge instructions address weight	180	90.6	0.81
HF-2	201	88.6	0.78
HF-3	116	94.0	0.88
HF-4	35	88.6	0.68

Table 2b.3.1. Reliability Findings by Williams et al, 2006.

Notes:

*HF-1 includes written instructions or educational material given to patient or caregiver at discharge or during the hospital stay addressing all of the following: activity level, diet, discharge medications, follow-up appointment, weight monitoring, and what to do if symptoms worsen.

2. 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% of hospitals exceeded the 80% 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-ofcare measures from July 2008-June 2009 with outcome-of-care measures from July 2006-June 2009. The second data set merges process-of-care measures from July 2007-June 2008 with outcome-of-care measures from July 2006-June 2009. Composite measures are calculated for hospitals where:

- 1. The hospitals reported rates for all four process and all two outcome-of-care indicators
- 2. Each process-of-care indicator had at least five cases and each outcome-of-care indicator had at least 25 cases.

The composite measures from these time periods were then compared. Across these two data collection periods, 2,906 hospitals had valid composite measures for HF.

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):

The Spearman correlation between composite measures computed in 2007-2008 and 2008-2009 was 0.41 (p<0.001), indicating moderate predictive validity of the composite. (See Table 2c.3.1) A large number of hospitals (around 55 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 (around 1 percent) occupy the first quartile in 2007-2008 and the fourth quartile in 2008-2009, and vice versa. Across the two separate time periods, around 36 percent of hospitals' categorizations differ by one quartile (i.e., 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.

	· ·				
		2008-2009	Reporting**		
2007-2008 Reporting*	Quartile 1	Quartile 2	Quartile 3	Quartile 4	Total
Quartile 1***	486	167	56	18	727
Quartile 2	183	310	186	47	726
Quartile 3	45	200	327	155	727
Quartile 4	13	49	158	506	726
Total	727	726	727	726	2,906
Spearman					
Correlation****	0.77				
	(0.00)				
Kappa Statistic	0.41				
	(0.00)				

Table 2c.3.1. Co	mparison of Con	nposite Measures,	by R	Reporting F	Period
------------------	-----------------	-------------------	------	-------------	--------

Notes:

* 2007-2008 reporting: process- and outcome-of-care measures for HF with a data collection period of July 2007 to June 2008

** 2008-2009 reporting: process- and outcome-of-care measures for HF with a data collection period of July 2008 to June 2009 *** Higher quartile categories indicate that the hospital had higher (i.e., better quality) composite measures.

**** P-values in parentheses.

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. We merged process-of-care indicators and outcome-of-care indicators for HF collected between July 2006 and June 2009. We estimated composite measures for 3,586 hospitals (out of potential 4,240 hospitals) for which:

- 1. The hospitals reported rates for all four process and all two outcome-of-care indicators
- 2. Each process-of-care indicator had at least five cases and each outcome-of-care indicator had at least 25 cases.

Background on Indicators Reported on Hospital Compare:

The indicators used in the construction of composites were drawn from *Hospital Compare*. The process-of-care indicators were drawn from Medicare hospital administrative claims data and medical record documents with discharge dates between July 2006 and June 2009. The hospital outcome-of-care indicators for 30-day risk-adjusted mortality and readmission for HF were based on Medicare claims for hospital stays with discharge dates between July 2006.

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 compare 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.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) :

Note: CMS has not decided how hospital performance will ultimately be displayed to consumers on Hospital Compare or to providers in hospital-specific reports. Table 2f.3.1 provides the number of hospitals in each of the three performance categories. These performance categories do not reflect how the composites will ultimately be displayed on Hospital Compare.

The total number of hospitals in each performance category is displayed in Table 2f.3.1. The table shows that there are meaningful differences in the overall composite score as 1,745 or around 48 percent of hospitals are categorized as being statistically different from the national average. Of the remaining 52 percent, around half of the hospitals' performances are significantly worse than the national average. The hospital performance category for the outcome-of-care domain is consistent with the hospital performance categories displayed on Hospital Compare for each of the indicators. That is, very few number of hospital s are in the "better than" or "worse than" the national rate categories.

	Categor	103	
	Pei	rformance Categ	ory
		No Different	
Type of	Worse than	than National	Better than
Composite	National Rate	Rate	National Rate
Overall	955	886	1,745
Process Domain	1,051	614	1,921
Outcome Domain	130	3,274	182

Table 2f.3.1. Number of Hospitals in Alternative Performance Categories



The distribution of composite scores by the following hospital characteristics:

- 1. Hospital bed size
 - 2. Ownership status

Μ

N

NA

- 3. Teaching status
- 4. Census region
- 5. Percentage of patients that was black.

Slight differences in the distribution were observed for hospital bed size, teaching status, census region, and race. Figures 2h.3.1-2h.3.4 present distributions for these characteristics. This analysis demonstrates that composite scores increase at most points along the distribution when hospital bed sizes increases as well as when the hospital is a teaching hospital (although teaching hospitals may also be more likely to be larger hospitals). This analysis also finds that there is very little difference in the distribution of the composite measure by the percentage of blacks served by hospital.





Figure 2h.3.3: Distribution of Composite Score, by Census Region



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



2e.4 If outcome or resource use measure is not risk adjusted, provide rationale:	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Scientific</i> Acceptability of Measure Properties?	2
Steering Committee: Overall, to what extent was the criterion, <i>Scientific Acceptability of Measure Properties</i> , 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.1 Current Use: 🗌 In use 🛛 Not in use	
3a.2 Use in a public reporting initiative (disclosure of performance results to the public at large) (<i>If used in a public reporting initiative, provide name of initiative(s), locations, Web page URL(s).</i> <u>If not publicly reported, state the plans to achieve public reporting within 3 years</u>):	
Following NQF endorsement, public reporting is expected on <i>Hospital Compare</i> sometime in 2012.	
3a.3 If used in other programs/initiatives (<i>If used in quality improvement or other programs/initiatives, name of initiative(s), locations, Web page URL(s).</i> <u><i>If not used for QI, state the plans to achieve use for QI within 3 years</i>):</u>	
Following NQF endorsement, CMS plans to publicly report this composite on <i>Hospital Compare</i> . CMS' current timetable calls for this public reporting to occur in 2012. CMS' experience indicates that hospitals closely scrutinize measures reported on <i>Hospital Compare</i> and consider these results as part of their quality improvement efforts.	
Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)	
3a.4 Data/sample (description of data/sample and size):	
Several studies suggest that the proposed composite measure will improve consumer understanding of hospital performance for HF patients, and be an asset to clinicians. In work that is directly relevant to the proposed measure, Borck et al held a series of focus groups that evaluated consumer and clinician understanding of condition-specific composite measures for AMI, HF, Pneumonia and SCIP that are very similar to the proposed measure. As well, their work evaluated understanding of AHRQ and Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) composite measures. In addition, work by Smith et al examined the interpretability of Hospital Compare data, including several of the component measures in the proposed composite. A further study by Peters et al also provides insight into consumer understanding of publicly reported hospital quality measures, while L&M Policy Research LLC specifically reports on consumer understanding of the 'readmissions' outcome-of-care indicator, one of two possible outcome-of-care indicators included in this composite.	
1. Borck, M, Thomas, C, & Gerteis, M. Transparency in Public Reporting: Consumer Testing and Enhancements to CMS's Compare Tools: <i>Topline Summary of Findings from Round #1 Interviews</i> <i>with Consumers,</i> April 9, 2009, and <i>Topline Summary of Findings from Round #2 Interviews with</i> <i>Consumers and Physicians, Composite measures of quality for Hospital Compare,</i> June 11, 2009. Memoranda to the Centers for Medicare & Medicaid Services.	3a C⊟
<i>Round 1</i> : Borck et al used a convenience sample of 21 consumers in the Baltimore, MD area. Participants ranged from 45-70 years old, were 67% women, and 48% Medicare beneficiaries. <i>Round 2</i> : Borck et al used a convenience sample of 18 consumers and 5 physicians from the Miami, FL	P M N

area. The group had an age range of 45 to 70 years old, and were made up of a majority of men and Medicare beneficiaries.

2. 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 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 out of 51 (92%) were over 65 years, and of the over 65 group, 53% were Medicare beneficiaries at risk for heart disease. Among the health care providers, 30% were nurses, 38% were primary care physicians, and the remainder were cardiologists and pulmonologists.

3. 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 employed a convenience sample of employed-age adults (18 - 64 years old, mean age of 37, 48% female, and 76% white) to determine whether providing only the most important quality information increase comprehension and information use. Half of the sample had lower levels of education (high school or less), 45% had health insurance and 74% had an annual household income of less than \$20,000.

4. 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, aged 50 - 70 years, most of whom were previously diagnosed with heart disease; 8 caregivers, aged 40 - 60 years; and 6 physicians who were primary care physicians, cardiologists, and pulmonologists.

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

1. 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 were utilized to assess comprehension of composite measures, organization and presentation of the site, and composite labels and descriptions.

2. 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 were first allowed to navigate the website independently and then asked a series of open-ended questions using an approved protocol during an approximately two-hour period.

3. 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 were presented with hospital quality data in a different format. In the first group, data on cost, quality, and non-quality information was unordered. In the second, cost and quality data was highlighted and presented first, while non-quality 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.

4. 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 them and their family members.

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

1. 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 conditionspecific composite measure. Most significantly, all respondents from Round 1 correctly interpreted the star ratings for the condition-specific composites (AMI, HF, Pneumonia and SCIP) and the HCAHPS composite measure. Round 1 also revealed that almost all participants preferred more descriptive definitions of the composites, and specifically that included a list of all the component measures making up the composite. Similarly to Round 1 findings, in Round 2 respondents were also found to be able to correctly interpret the star ratings for condition-specific quality ratings composites and the HCAHPS composite. However, some respondents in Round 2 did not understand that the conditionspecific 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, should also be easy for consumers to use. Moreover, any composite definition posted on Hospital Compare should include a list of all component measures.

2. 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 is intended to address this very issue by creating a single benchmark that enables consumers to evaluate the quality of care at a given hospital for a given condition.

3. 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.

Similarly to Smith et al (2005), Peters et al (2007) determined that less is more with regards 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% of respondents choose the highest quality hospital Y when only

	<i>r</i> .
the quality information was shown, while in the other two formats it was by selected 48% (ordered group) and 40% (unordered group). Such results reinforce the idea that a composite measure may enhance the utility of hospital quality data for consumers.	
4. L&M Policy Research LLC. Report to the Centers for Medicare & Medicaid Services: <i>Recommendations for Incorporating Hospital Readmission Data into the Hospital Compare Website</i> . 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. While 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 <u>NQF-endorsed measures</u> to components and/or composite	
3b.1 NQF # and Title of similar or related measures:	
All components of this composite measure are all NQF-endorsed. However there are currently no NQF-endorsed composite measures that provide a single indication of a hospital's quality of care for HF 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:	
1.NQF #0531Patient Safety for Selected Indicators (AHRQ)Endorsed June 19, 20092.NQF #0530Mortality for Selected Conditions (AHRQ)Endorsed June 19, 2009	
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?	3b
 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 M N N
 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. 3c. Distinctive or Additive Value 3c.1 Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: 	3b C P M NA
 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. 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 HF patients. 	3b C P M NA
 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. 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 HF patients. 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: 	3b C P M N N N N N C C P
 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. 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 HF patients. 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. 	3b C P M N N N N N C P M M N
 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. 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 HF patients. 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. 3d. Decomposition of Composite 3d.1 Describe the information that is available from decomposing the composite into its components: 	3b C P M N N N N N N 2d
 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. 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 HF patients. 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. 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: 	3b C P M N N N N N N S C P M N N S d C P

NQF Review #	ŧ:
 Percent of HF Patients Given ACE Inhibitor or ARB for LVSD Percent of HF Patients Given Smoking Cessation Advice/Counseling Heart Failure (HF) 30-day Mortality Heart Failure (HF) 30-day Readmission 	
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 HF 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 HF, achieving the purpose of the composite.	3e C P M N
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, <i>Usability</i> , 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 <u>all</u> the data elements that are needed to compute measure scores generated? (<i>Check all that apply</i>) Data are generated as a byproduct of care processes <u>during</u> care delivery (<i>Data are generated and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition</i>) Coding/abstraction performed by someone other than person obtaining original information (<i>e.g., DRG, ICD-9 codes on claims; chart abstraction for quality measure, registry</i>) Survey Other (<i>e.g., patient experience of care surveys, provider surveys, observation</i>), <i>Please describe</i>: 	4a C P M N
 4b. Electronic Sources 4b.1 Are <u>all</u> the data elements available electronically? (<i>elements that are needed to compute measure scores are in defined, computer-readable fields, e.g., electronic health record, electronic claims</i>) ∑ Yes ∑ No 4b.2 If no, specify the near-term path to achieve electronic capture by most providers. N/A 	4b C P
date	
 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 — M — N
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:	4e C P M N

	<i>t</i> .
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 HF 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/A4e.4 Business case documentation: N/A	
If the component measures are <u>combined at the patient level</u> , complete 4c.	4c H□
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	
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Feasibility</i> ?	4
Steering Committee: Overall, to what extent was the criterion, <i>Feasibility</i> , met? Rationale:	4 C P M
	N
RECOMMENDATION	N
RECOMMENDATION Steering Committee: Do you recommend for endorsement? Comments:	N Y N A
RECOMMENDATION Steering Committee: Do you recommend for endorsement? Comments: CONTACT INFORMATION	N N A
RECOMMENDATION Steering Committee: Do you recommend for endorsement? Comments: CONTACT INFORMATION Cont 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	
RECOMMENDATION Steering Committee: Do you recommend for endorsement? Comments: CONTACT INFORMATION Cont 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:	N N A
RECOMMENDATION Steering Committee: Do you recommend for endorsement? Comments: CONTACT INFORMATION Content of Conters 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 State: MA ZIP: 02139	
RECOMMENDATION Steering Committee: Do you recommend for endorsement? Comments: 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: Mathematica Policy Research 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:	
RECOMMENDATION Steering Committee: Do you recommend for endorsement? Comments: CONTACT INFORMATION 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 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: Mathematica Policy Research I Measure Steward Measure Developer 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: Mathematica Policy Research I Measure Steward Measure Developer First Name: Marian Last Name: Wrob	
RECOMMENDATION Steering Committee: Do you recommend for endorsement? Comments: CONTACT INFORMATION 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: Mathematica Policy Research 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: Mathematica Policy Research I Measure Steward I Measure Developer First Name: Marian Last Name: Wrobel Credentials (MD, MPH, etc.): Ph.D. Email: MWrobel@mathematica-mpr.com Telephone: 617-301-8971 ext: Co.6 List any additional organizations that sponsored/participated in measure development:	
RECOMMENDATION Steering Committee: Do you recommend for endorsement? CONTACT INFORMATION CONTACT INFORMATION Contact: INFORMATION Contact INFORMATION Contact INFORMATION Contact Information Property Owner) Organization: Centers for Medicare & Medicald 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: Mathematica Policy Research Street Address: Mathematica Policy Research Street Address: Mathematica Policy Research Street Address: Mathematica Policy Research Street Address: Mathematica Policy Research Street Address: Mathematica Policy Research Street Address: Mathematica Policy Research Street Address: Mathematica Policy Research Street Address: Mathematica Policy Research	

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*): Initial: 12/13/10 Resubmission: 3/15/11

MEMORANDUM

Policy Research, Inc. 955 Massachusetts Avenue, Suite 801 Cambridge, MA 02139

MATHEMAT

Cambridge, MA 02139 Telephone (617) 491-7900 Fax (617) 491-8044 www.mathematica-mpr.com

TO:	National Quality Forum	
FROM:	Marian Wrobel, Mai Hubbard, Jessica Ross, Robert Schmitz	DATE: 3/15/2011
SUBJECT:	Heart Failure Composite Measure	

We have attached the submission form supporting the Centers for Medicare and Medicaid Services' (CMS) composite measure of inpatient quality for heart failure (HF). This submission revises material we provided to NQF on December 14, 2010. This memorandum first describes the composite in brief and then outlines the ways in which the material reported in the form differs from the earlier version.

COMPOSITE SUMMARY

1. Importance

HF is one of the most costly diseases for the Medicare program. More than 230,000 beneficiaries were hospitalized for heart failure in 2009. Medicare payments for these hospital stays were nearly \$2 billion. Moreover, hospital patients with HF have one of the highest rates of hospital readmission. Among hospitalized patients studied by Jencks et al. (2009), HF patients had the highest rate of rehospitalization at nearly 27 percent.

CMS provides consumers with indicators of hospital quality of care on its *Hospital Compare* website. At present, it reports four NQF-endorsed process-of-care indicators and two endorsed outcome-of-care indicators for HF. CMS seeks to provide consumers with a composite of these indicators, summarizing all six. The composite described here is formative in nature, as it seeks to summarize existing indicators rather than to reflect an underlying latent quantity. The measure provides to consumers a single useful summary of the quality of HF care.

2. Properties

The proposed composite measure is computed as the simple sum of process and outcome domain scores at the hospital level, $S_{p,i}$ and $S_{o,i}$. Each domain score is computed as the weighted sum of the ratio of actual to expected scores. Thus the process domain score for hospital *i* is computed as:

$$S_{P,i} = w_1 \frac{g_{1,i}}{\overline{g}_1} + w_2 \frac{g_{2,i}}{\overline{g}_2} + w_3 \frac{g_{3,i}}{\overline{g}_3} + w_4 \frac{g_{4,i}}{\overline{g}_4},$$

where the g_j are the four process indicators, the \overline{g}_k are the national means of the indicators, and the w_k are the share of the number of observations nationally for each process indicator. That is,

An Affirmative Action/Equal Opportunity Employer

MEMO TO: National Quality Forum
FROM: Marian Wrobel, Mai Hubbard, Jessica Ross, Robert Schmitz
DATE: 3/15/2011
PAGE: 2

the indicators are "denominator weighted" in each domain score. The weight for each indicator is thus proportional to the number of patients who are at risk of not receiving the care described by that indicator. We recognize the appeal of weighting component indicators by their clinical importance, but are aware, as NQF itself has noted, that consensus on such weights is difficult to achieve. We have adopted denominator weights as a transparent and recognized approach.

The outcome domain score for hospital *i* is defined similarly:

$$S_{O,i} = v_1 \frac{h_{1,i}}{\overline{h_1}} + v_2 \frac{h_{2,i}}{\overline{h_2}},$$

where the h_i are outcome indicators and v_k are the share of the number of observations nationally for each outcome indicator. The HF composite for hospital *i* is the simple average of the domain scores:

HF Composite_i
$$= \frac{1}{2} (S_{P,i} + S_{O,i})$$

Values of composite scores are strongly consistent from year to year. The Spearman correlation coefficient between values in 2008 and 2009 was 0.77. Although we have not argued for reflective properties of the composite, exploratory factor analysis does indicate a single underlying construct.

3. Usability

The supporting statement reviews studies that indicate that the proposed measure will improve consumer understanding of hospital performance. CMS currently plans to report HF composite values on the *Hospital Compare* website in 2012. It will always be reported together with the underlying process and outcome indicators. The composite will be useful to hospitals as well, though hospitals will surely seek to examine values of its component indicators to understand how its value was determined.

4. Feasibility

The composite is feasible to construct as it is computed entirely from Medicare claims and quality data currently reported to CMS. The measure places no additional reporting burden on hospitals.

REVISIONS TO THE FORM

We have revised material in the form to respond to issues raised in the committee's discussion of the proposed AMI composite measure and also in the communication from NQF to

MEMO TO: National Quality Forum
FROM: Marian Wrobel, Mai Hubbard, Jessica Ross, Robert Schmitz
DATE: 3/15/2011
PAGE: 3

measure developers on February 23. Changes to the measure and documentation are described below:

- **Treatment of disparities has been expanded.** The committee has expressed a particular interest in variation in measured quality across geographic areas, type of hospital, and racial/ethnic composition of the patient population. To address these concerns, we have provided additional information describing differences in mean composite scores by hospital size, teaching status, region, and the proportion of nonwhites in total discharges. These results are shown in response to item 2h of the submission.
- The minimum sample size for process indicators has been increased. The measure presented here considers constituent process indicators to be non-missing if they are computed on a sample of five or more patients. The earlier proposed measure required a minimum of just one patient. The new minimum ensures that the maximum standard error for any process indicator is 0.22.
- Values for all constituent indicators are now required. The revised measure requires that all constituent outcome and process indicators be present for a hospital. The earlier proposed measure required that one outcome indicator and two process indicators be present. The change eliminates the need to impute values for missing indicators.
- We have changed and simplified the way in which the individual indicators and the final composite are standardized. First, each individual indicator is standardized to one by dividing by the mean. The indicators are weighted and combined in such a way that the final composite remains standardized to one. This change leads to a composite that is centered on one and has an inter-quartile range of 0.074 (1.047-0.973) and corrects a problem in which the earlier composite appeared extremely tightly distributed.

The changes described above produce a measure that is better documented, more reliable, and easier to understand than the measure described in our earlier submission. Eliminating imputation and requiring each hospital to report process indicators based on a sample of five rather than one case, resulted in a reduction of about 8 percent in the number of hospitals for which the composite measure was calculated – from 3,900 to 3,586. The properties of the measure, however, differ little from those appearing in the previous submission.

The National Quality Forum Composite Measure of Hospital Quality for HF

Appendix A Technical Supplement

Submitted By:

Mathematica Policy Research, Inc Marian Wrobel, Ph.D., Project Director Bob Schmitz, Ph.D. Mai Hubbard, Ph.D. Jessica Ross, M.P.H.

> Boston University Jim Burgess, Ph.D. Gary Young, Ph.D.

Prepared for: Centers for Medicare and Medicaid Services

March 2011

TABLE OF CONTENTS

Section 1	Background	2
1.1	Overview	3
Section 2	Method of Scoring and Aggregation	5
2.1	Estimation of the Composite Measure	6
2.2	Estimation of Reliability-Weight-Adjusted Measures	7
Section 3	Performance Discrimination	9
3.1	Method of Discriminating Performance	10
Section 4	Results	12
4.1	Results for Section 2k.3	13
4.2	Results for Section 2k.5	13
4.3	Results for Section 21.3	14
4.4	Results for Section 2h.2	15

SECTION 1 BACKGROUND

1.1 Overview

The composite measure of quality of hospital care for HF 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 four individual process-of-care indicators and two outcome-of-care indicators meant to capture the quality of hospital care provided to patients with HF. 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:

- 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: Acute Myocardial Infarction (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, we present the mapping between CMS' policy goals and methodological decisions in tabular form.

Table 1d.2.1.	CMS	Policy	Goals	for	Composite	Measures	and	Associated	Methodolog	gical
Decisions										

Policy Goals	Methodological Decisions
Summarize measures on Hospital Compare in a single, useful, condition- specific composite	• Include the same set of process and outcome measures as Hospital Compare
Produce differences in composite values that are clinically and statistically meaningful and reflect true differences in underlying quality	 Enter component measures as values, not ranks, so that slight differences in measured performance do not potentially lead to large differences in the composite value for topped-off measures For process measures, adjust component measures for reliability so that random variation does not drive small hospitals to extremes
Results available for a large number of hospitals	• Process measures are available when the number of eligible discharges is five or more; outcome variables are available when the number of eligible discharges is 25 or more
Focus more on measures relevant to more patients	• Construct process and outcome composites using weights based on national denominators
Method is scientifically acceptable and acceptable to stakeholders	 Adopt an approach that is similar to that used for AHRQ quality indicators (QIs) Note: AHRQ QIs are NQF-endorsed and widely reported
Method accommodates changes in the set of measures on Hospital Compare Method can be used for multiple conditions Relative weighting of process and outcome domains does change when measures are added to or deleted from one domain	 Method is based on general principles, not on the specific statistical performance of a group of measures Process and outcome domains are statistically standardized before they are added together
Method is relatively simple Hospitals can duplicate results	 Use equal weighting to combine process and outcome domains Reliability weights are a function of a hospital's number of cases and national parameters

SECTION 2 METHOD OF SCORING AND AGGREGATION

2.1 Estimation of the 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 two domain scores related to hospital inpatient quality. The first domain is comprised of four process-of-care indicators and the second domain is comprised of two outcome-of-care indicators. All of these indicators are publically reported by the CMS on *Hospital Compare* and NQF endorsed.

To construct the process-of-care domain, the process-of-care indicators were 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). More information regarding the reliability-weight adjustment is available in Section 2.2. Then, each indicator was standardized by dividing by the national mean of the indicator. Since the outcome-of-care indicators have already been risk-standardized using a hierarchical generalized linear modeling technique, the outcome-of-care indicators, the outcome-of-care indicators were not reliability-weight adjusted. Similarly to the process-of-care indicators, the outcome-of-care indicators were also was standardized by dividing by the national mean of the indicator.

Consistent with the approach used for the AHRQ measures, CMS used denominator weighting in constructing the process- and outcome-of-care domains. Denominator weighting places relatively more weight on measures that apply to relatively more patients nationally. More specifically, the process of care domain for hospital j = 1, ..., J can be described as a denominator weighted average of a standardized reliability-weight adjusted process-of-care indicator k=1,...K,

$$P_j^* = \sum_{k=1}^{K} \left(\frac{\sum_{j=1}^{J} n_{jk}}{\sum_{k=1}^{K} \sum_{j=1}^{J} n_{jk}} * \frac{P_{jk}^*}{P_k^{nat}} \right)$$
(eq. 2.1.1)

where P_k^{nat} is the national rate of a process-of-care indicator and n_{jk} is the total number of cases for a process-of-care indicator at hospital *j*.

Similarly, the outcome-of-care sub-composite score is estimated used denominator weighting. That is

$$O_{j}^{*} = \sum_{l=1}^{L} \left(\frac{\sum_{j=1}^{J} n_{jl}}{\sum_{l=1}^{L} \sum_{j=1}^{J} n_{jl}} * \frac{O_{jl}^{*}}{O_{l}^{nat}} \right)$$
(eq. 2.1.2)

where n_{jl} is the number of hospital cases for HF outcome-of-care indicator l=1...,L, in hospital j=1,...,J and O_{jl}^* is the risk-standardized outcome-of-care score.

The overall composite score (C_i^*) is then estimated as a simple average of the two domains:

$$C_j^* = \frac{1}{2} (P_j^*) + \frac{1}{2} (O_j^*)$$
(eq. 2.1.3)

2.2 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 (n_{jk}) increases, the weight approaches one.

First, let:

Signal variance
Within variance
Hospital-specific rate for process-of-care indicator k
National rate for process-of-care indicator k
Total number of cases in hospital j for indicator k
Total number of hospitals for indicator k
Process-of-care indicator
Hospital index

Then the reliability-weight adjusted estimator (P_{jk}^*) is

$$P_{jk}^* = W_{jk}P_{jk} + (1 - W_{jk})P_k^n$$
(eq. 2.2.1)

where W_{jk} is the reliability-weight:

$$W_{jk} = \frac{\sigma_{sk}^2}{\sigma_{sk}^2 + \sigma_{wjk}^2}$$
(eq. 2.2.2)

 σ_{sk}^2 is the signal variance:

$$\sigma_{sk}^{2} = \frac{\sum_{i=1}^{J} (P_{ik} - P_{k}^{n})^{2}}{N_{k}} - \frac{\sum_{i=1}^{J} P_{ik} (1 - P_{ik})}{\sum_{i=1}^{J} n_{ik}}$$
(eq. 2.2.3)

and σ_{wjk}^2 is the within variance:

$$\sigma_{wjk}^{2} = \frac{\sum_{i=1}^{J} P_{ik} (1 - P_{ik}) \frac{n_{ik}}{\sum_{l=1}^{J} n_{lk}}}{n_{jk}}$$
(eq. 2.2.4)

SECTION 3 PERFORMANCE DISCRIMINATION

3.1 Method for Discriminating Performance

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.

More specifically, the standard error for a specific hospital is calculated as follows. First, we let:

P_{jk}^*	Hospital-specific reliability-weight-adjusted rate for process-of-care
	indicator k
O_{jl}^*	Risk-standardized hospital-specific rate for process-of-care indicator l
n_{jk}	Total number of cases in hospital j for indicator k
N _k	Total number of hospitals for indicator k
μ_P	Mean of process domain composite
μ_O	Mean of outcome domain composite
σ_P	Standard deviation of process domain composite
σ_0	Standard deviation of outcome domain composite
$k = 1, \dots K$	Process-of-care indicator
l = 1, L	Outcome-of-care indicator
j = 1,, J	Hospital index

The hospital's process-of-care domain composite score (P_j^*) is estimated as a denominator weighted average of the standardized reliability-weight-adjusted process-of-care indicator rates:

$$P_j^* = \sum_{k=1}^{K} \left(\frac{\sum_{j=1}^{J} n_{jk}}{\sum_{k=1}^{K} \sum_{j=1}^{J} n_{jk}} * \frac{P_{jk}^*}{P_k^{nat}} \right)$$
(eq. 2.3.1)

The hospital's outcome-of-care domain composite score (O_j^*) is estimated as a denominator weighted average of the standardized risk-adjusted outcome-of-care indicator rates:

Composite Measure of Hospital Quality for Heart Failure Appendix A March 2011

$$O_{j}^{*} = \sum_{l=1}^{L} \left(\frac{\sum_{j=1}^{J} n_{jl}}{\sum_{l=1}^{L} \sum_{j=1}^{J} n_{jl}} * \frac{O_{jl}^{*}}{O_{l}^{nat}} \right)$$
(eq. 2.3.2)

The composite measure (C_j) is a simple average of the normalized process-of-care and outcome-of-care sub-composites.

$$C_j^* = \frac{1}{2} (P_j^*) + \frac{1}{2} (O_j^*)$$
(eq. 2.3.3)

Therefore, the variance of the composite measure $Var(C_j)$ can be estimated as

$$\begin{aligned} \operatorname{Var}(\mathcal{C}_{j}^{*}) &= \operatorname{Var}\left[\frac{1}{2}P_{j}^{*} + \frac{1}{2}O_{j}^{*}\right] \\ &= \left(\frac{1}{2}\right)^{2}\operatorname{Var}\left[\sum_{k=1}^{K}\left(\frac{\sum_{j=1}^{J}n_{jk}}{\sum_{k=1}^{K}\sum_{j=1}^{J}n_{jk}}\right)\frac{P_{jk}^{*}}{\mu_{Pk}^{*}} + \sum_{l=1}^{L}\left(\frac{\sum_{l=1}^{J}\sum_{j=1}^{J}n_{jl}}{\sum_{l=1}^{L}\sum_{j=1}^{J}n_{jl}}\right)\frac{O_{jl}^{*}}{\mu_{Ol}^{*}}\right] \\ &= \left(\frac{1}{2}\right)^{2}\left\{\frac{1}{\left(\sum_{k=1}^{K}\sum_{j=1}^{J}n_{jk}\right)^{2}}\sum_{k=1}^{K}\left[\left(\frac{\sum_{j=1}^{J}n_{jk}}{\mu_{Pk}^{*}}\right)^{2}\frac{P_{jk}^{*}(1-P_{jk}^{*})}{n_{jk}}\right] \\ &+ \frac{1}{\left(\sum_{l=1}^{L}\sum_{j=1}^{J}n_{jl}\right)^{2}}\sum_{l=1}^{L}\left[\left(\frac{\sum_{j=1}^{J}n_{jl}}{\mu_{Ol}^{*}}\right)^{2}\operatorname{Var}(O_{jl}^{*})\right]\right\} \end{aligned}$$
(eq. 3.4)

given the following assumptions:

A1.
$$\sigma_P$$
, μ_P and σ_O , μ_O are constants
A2. $\operatorname{cov}(P_{jm}^*, P_{jn}^*) = 0 \quad \forall \ m \neq n$
A3. $\operatorname{cov}(O_{jm}^*, O_{jn}^*) = 0 \quad \forall \ m \neq n$

A4. $\operatorname{cov}(P_{jm}^*, O_{jn}^*) = 0$

SECTION 4 RESULTS

4.1 Results for Section 2k.3

Table 2k.3.1. Comparison of Distribution of HF Composite Measure by Weighting Method

	Equal	Differential
Percentile	Weighting	Weighting
Min	0.65	0.64
1%	0.77	0.75
5%	0.87	0.86
10%	0.92	0.91
25%	0.98	0.97
50%	1.02	1.02
75%	1.04	1.05
90%	1.06	1.07
95%	1.07	1.08
99%	1.08	1.09
Max	1.10	1.12
Mean	1.00	1.00
Ν	3,586	3,586

4.2 Results for Section 2k.5

Table 2k..1. Comparison of Distribution of HFComposite Measure by Scoring Method

		Absolute Scoring
	Absolute Scoring	Index with
	Index with	Reliablity
	Reliability	Weights (Old
Percentile	Weights	Version)
Min	0.64	71.02
1%	0.75	75.75
5%	0.86	78.54
10%	0.91	79.64
25%	0.97	81.09
50%	1.02	82.21
75%	1.05	83.11
90%	1.07	83.83
95%	1.08	84.24
99%	1.09	85.07
Max	1.12	86.86
Mean	1.00	81.91
Ν	3,586	3867

4.3 Results for Section 21.3

	HF	1*	HF	2*	HF3*		HF4*		Survival**		Readmission**	
	All	Included	All	Included	All	Included	All	Included	All	Included	All	Included
Percentile	Hospitals	Hospitals	Hospitals	Hospitals								
Min	0.00	9.09	0.00	0.00	0.00	0.00	0.00	9.13	6.60	6.60	17.30	17.30
1%	24.88	50.00	0.00	13.19	0.00	30.00	2.42	37.04	8.00	7.90	20.40	20.30
5%	60.20	68.56	18.00	39.81	40.00	61.69	43.14	64.06	8.90	8.90	21.70	21.60
10%	71.57	75.29	37.76	51.94	64.29	75.31	62.93	76.84	9.40	9.40	22.40	22.40
25%	82.35	83.33	61.93	66.29	86.54	89.52	84.57	89.42	10.30	10.20	23.50	23.50
50%	89.75	89.80	77.42	79.21	96.69	96.92	94.73	95.58	11.20	11.20	24.60	24.60
75%	94.81	94.39	87.54	88.10	100.00	99.75	98.03	98.26	12.20	12.20	25.90	26.00
90%	98.71	97.46	94.45	94.66	100.00	100.00	99.33	99.33	13.20	13.20	27.20	27.30
95%	100.00	99.14	97.30	97.24	100.00	100.00	99.72	99.69	13.90	14.00	28.10	28.20
99%	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	15.30	15.30	29.90	30.00
Max	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	18.20	17.30	32.40	32.40
Mean	86.31	87.49	71.21	75.26	87.85	91.32	86.90	90.98	11.28	11.24	24.74	24.75
Ν	4,104	3,586	4,182	3,586	4,087	3,586	4,198	3,586	3,890	3,586	3,937	3,586

Table 21.3.1: Comparison of Hospitals' Rates for Hospitals the Full Sample and for Hospitals Included in the Composite Caluclation

Notes:

* HF1: Percent of HF Patients that Received Discharge Instructions; HF2: Percent of HF Patients with Evaluation of LVS Function; HF3: Percent of HF Patients Given ACE Inhibitor or ARB for LVSD; HF4: Percent of HF Patients Given Smoking Cessation Advice/Counseling.

** Survival: 30-day risk-adjusted survival rate; Readmission: 30-day risk-adjusted lack of readmission.

4.4 Results for Section 2h.2

	Bed Size						
Percentile	0-49	50-199	200-399	400+			
Min	0.64	0.68	0.74	0.83			
1%	0.69	0.77	0.90	0.90			
5%	0.76	0.87	0.95	0.95			
10%	0.80	0.92	0.97	0.97			
25%	0.90	0.97	1.01	1.00			
50%	0.97	1.01	1.03	1.03			
75%	1.02	1.05	1.05	1.06			
90%	1.05	1.07	1.07	1.07			
95%	1.06	1.08	1.08	1.08			
99%	1.08	1.09	1.09	1.09			
Max	1.09	1.12	1.11	1.10			
Mean	0.95	1.00	1.03	1.03			
Ν	664	1,539	868	437			

Table 2h.2.1. Comparison of Distribution of Composite Measure, by Bed

 Size

Table 2h.2.2.	Comparison	of Distribution of	of Composite
Measure, by O	wnership Ty	/pe	

	Ownership					
Percentile	Government	Not for Profit	For Profit			
Min	0.64	0.64	0.64			
1%	0.70	0.78	0.75			
5%	0.77	0.89	0.86			
10%	0.83	0.93	0.92			
25%	0.93	0.98	0.98			
50%	0.99	1.02	1.02			
75%	1.03	1.05	1.05			
90%	1.05	1.07	1.07			
95%	1.07	1.08	1.08			
99%	1.08	1.09	1.10			
Max	1.10	1.11	1.12			
Mean	0.96	1.01	1.00			
N	659	2,257	592			

	Teaching Hospital			
Percentile	Yes	No		
Min	0.88	0.64		
1%	0.90	0.74		
5%	0.97	0.85		
10%	0.98	0.91		
25%	1.01	0.97		
50%	1.03	1.01		
75%	1.06	1.05		
90%	1.07	1.07		
95%	1.08	1.08		
99%	1.09	1.09		
Max	1.10	1.12		
Mean	1.03	1.00		
N	268	3,240		

Table 2h.2.3. Comparison of Distribution of Composite Measure, by Teaching Hospital Status

Table 2h.2.4. Comparison of Distribution of Composite Measure, by CensusRegion

	Census Region			
Percentile	Northeast	South	Midwest	West
Min	0.72	0.64	0.71	0.68
1%	0.77	0.73	0.75	0.74
5%	0.93	0.83	0.86	0.86
10%	0.97	0.89	0.91	0.91
25%	1.00	0.96	0.98	0.96
50%	1.03	1.01	1.02	1.01
75%	1.05	1.04	1.05	1.05
90%	1.07	1.07	1.07	1.07
95%	1.08	1.08	1.08	1.08
99%	1.09	1.09	1.09	1.09
Max	1.11	1.12	1.10	1.11
Mean	1.02	0.99	1.00	1.00
Ν	547	1,424	920	587

	Percentage of Black Patients*				
Percentile	0	>0 and ≤15	>15 and ≤30	>30	
Min	0.64	0.64	0.69	0.69	
1%	0.72	0.76	0.74	0.75	
5%	0.78	0.88	0.87	0.86	
10%	0.85	0.93	0.92	0.91	
25%	0.93	0.98	0.98	0.98	
50%	0.99	1.02	1.02	1.02	
75%	1.03	1.05	1.05	1.04	
90%	1.06	1.07	1.07	1.06	
95%	1.07	1.08	1.08	1.07	
99%	1.08	1.09	1.09	1.09	
Max	1.09	1.11	1.12	1.09	
Mean	0.97	1.01	1.00	1.00	
Ν	602	1,960	488	536	

Table 2h.2.5. Comparison of Distribution of Composite Measure, by Percentage of Patients that are Black

The National Quality Forum Composite Measure of Hospital Quality for HF

Appendix B

Submitted By:

Boston University Jim Burgess, Ph.D. Gary Young, Ph.D.

Prepared for: Centers for Medicare and Medicaid Services

March 2011

BACKGROUND

Composite measures are used in many contexts or settings to provide a broad picture of the performance, behavior, traits and other characteristics of individuals or other types of entities. In general, composite measures combine quantitatively two or more separate measures into a single measure or index. Within health care, a composite measure can be formed by combining quantitatively the performance data of providers across multiple measures.

Such composite measures of provider performance serve two primary goals. First it summarizes a large amount of information about the performance of a provider. This type of summary can be useful for giving consumers provider-related performance information. Much research has shown that consumers find it difficult and frustrating to sort through multiple performance measures to arrive at a conclusion regarding the performance of a provider from whom they are contemplating receiving care (Hibbard et al., 2000; Hibbard, 2001). Thus composites are a potentially useful tool for sponsors of consumer report cards and other types of vehicles for disseminating information about provider performance to consumers. Providers also may benefit when their performance information is presented in a summary form if the summary offers insight about opportunities for improvement.

Second, it increases measurement reliability for providers. As provider profiling and consumer report cards have become widely used, researchers have raised concerns about the reliability of performance measurement. Studies have demonstrated that measurement reliability is often below acceptable levels because of small sample sizes for providers (Zaslavsky, 2001). The construction of composites may be used to address this problem by combining, for a given provider, the number of patients across the multiple measures.

With respect to the information summarized, composites for healthcare measures are likely to comprise process measures, outcome measures or some combination of the two. Although in the field of health services research, process measures are sometimes treated as an intermediate measure for outcomes within conceptual models of quality of care, there is no consensus that process measures are not important in their own right for assessing quality of care. First, it is not clear that process scores consistently correspond with outcomes as studies examining the statistical correlations between process and outcome measures often report mixed results. In addition, more recent studies using sophisticated measurement techniques seem to indicate that they are not related strongly (e.g. Jha et al., 2007; Ryan et al., 2009). Second, for quality improvement, processes always are much more under the control of providers than are outcomes as they offer guidance as to what actions provider can undertake to improve scores. As such, many providers appear to value process measures for purposes of quality assessment.

There are two general approaches for constructing composites (Shwartz et al., 2009). One approach is to construct "reflective" composites. A reflective composite seeks to combine multiple measures that theoretically are believed to be linked to an underlying construct that cannot be directly measured such as quality or intelligence. The construction of a reflective construct requires that the individual measures be highly correlated as they are treated theoretically as representing different dimensions of the same construct. The other approach is to construct "formative" composites. A formative composite is essentially a combination of

multiple measures that are intended to provide useful summary information but without a strong theoretical rationale that they are linked to the same construct. As such, there is no expectation that the individual measures comprising the composite will be highly correlated or meet other psychometric tests that are considered standard for the construction of a valid reflective composite. In particular, then, reflective measures may gain validity and reliability by summarizing information from individual indicators in a condensed form. Such a result may or may not hold for particular formative measures.

CMS HOSPITAL COMPARE COMPOSITES

CMS has developed composite measures for four conditions that are part of the accepted set of measures from the CMS Hospital Compare system: Acute Myocardial Infarction (AMI), Heart Failure (HF), Pneumonia (PN), and Surgical Care Improvement Project (SCIP). For three of these four conditions (i.e., AMI, HF, and PN), both process and outcome measures are available for constructing composites. For SCIP, process measures are available only. For constructing the composites, the process and outcome measures were treated as separate domains. All the measures comprising the composites have previously been reviewed and endorsed by the National Quality Forum (NQF). Because CMS plans to include these composite measures in the Hospital Compare website, which is a consumer-oriented tool for comparing provider performance, a primary goal is to summarize information in a way that will be helpful to consumers.

The construction of these composites was conducted in manner that is consistent with a formative approach. There are several considerations that are relevant to this decision. First, the process by which the measures comprising each composite evolved and were chosen for Hospital Compare did not take place with a reflective construct in mind. The measures were developed, evaluated, and considered for NQF endorsement separately, each on their own merits. Thus, we consider these constructs formative in that they summarize an array of measures for that condition. Second, each of the four conditions is complex in etiology and treatment, so that it is difficult or even impossible to condense the measures into simple and valid conceptual constructs as would be seen in reflective composites. Yet, the decisions from a patient, provider, and healthcare system level on evaluating quality for individual treatment conditions need to be made. We cannot pick and choose to take the treatment of one hospital for one measure and another hospital for another measure; the treatment comes as a package. Third, composites are intended to be flexible for future additions or deletions of measures. CMS policy on the appropriate measures for these conditions and possibilities for additional conditions will adapt to measure development opportunities and changes in the evidence base underlying both process and outcome measures over time. Finally, the process and outcome measures themselves have different theoretical constructs, are affected differently by the actions of providers, and may not be causally related to each other. As such, for each of these four conditions now, and for any new conditions that are added, formative composites can be developed following the technical procedures that have been outlined in the initial NQF submissions for each of these composites.

A key technical decision as to the construction of the composites was to weight the process and outcome domains equally by standardizing each domain score, before combining into a single composite score. The decision to weight equally was based on the consideration that no strong theoretical foundation existed for assigning differential weights. In this sense, the rationale is similar to the decision to construct the composites as a formative measure. Since the measures are not necessarily drawn from a consistent unifying underlying construct, there may not really be a population standard deviation for each measure to be estimating by the sample standard deviation. Also, for true equal weighting to be achieved, standardization of the domain scores is necessary. This is because the impact of any measure on a composite with equal weighting will be proportional to the standard deviation of the underlying measure. Measures which vary more will have greater influence on the composite measure and the ranking of entities measured. Zscore methods to normalize measures to mean 0 and standard deviation of 1 are possible to equalize the influence across all measures, but this is undesirable since it greatly inflates the influence of measures with very small standard deviation measured differences that likely have little to no clinical or practical significance. In fact, for practical implementation of a composite measure where expert opinion is not being brought to bear on weighting, equal weighting where the standard deviation impact is allowed to pass through to the composite measure actually is more acceptable.

REFERENCES

Hibbard, JH (2001). Being Strategic with Public Reports. *The International Journal for Quality in Health Care*, Vol. 13, No. 5, pp. 353-355.

Hibbard, JH, Harris-Kojetin L, Mullin P, Lubalin JS, Garfinkel SA (2000). Increasing the Impact of Health Plan Report Cards by Addressing Consumer Concerns. *Health Affairs*, Vol. 19, No. 5, pp. 138-143.

Jha, Ashish K., E. John Orav, Zhonghe Li and Arnold M. Epstein (2007). The Inverse Relationship Between Mortality Rates And Performance In The Hospital Quality Alliance Measures, *Health Affairs*, Vol. 26, No. 4, pp. 1104-1110.

Ryan, Andrew M., James F. Burgess, Jr., Christopher P. Tompkins, Stanley S. Wallack (2009). The Relationship between Medicare's Process of Care Quality Measures and Mortality. *Inquiry*, Vol. 46, No. 3, pp. 274-290.

Shwartz, Michael, James F. Burgess, Jr., and Dan Berlowitz (2009). Benefit-of-the-Doubt Approaches for Calculating a Composite Measure of Quality, Health Services and Outcomes Research Methodology, Vol. 9, No. 4, pp. 234-251.

Zaslavsky, Alan M. (2001). Statistical issues in reporting quality data: small samples and casemix variation, *International Journal for Quality in Health Care*, Volume 13, Issue 6, pp. 481-488.