

THE NATIONAL QUALITY FORUM

Hospital Outcomes & Efficiency
 Technical Advisory Panel - February 2009
 Steering Committee - March 2009

Summary of Review of Measures

NQF Evaluation Criteria: **I**=Importance to measure and report; **S**=Scientific acceptability of measure properties; **U**=Usability; **F**=Feasibility
Importance to measure and report: this is a threshold criterion and the Committee votes: **Y**=yes, **N**=no, or **A**=abstain. Measures that do not pass the importance criterion are not further evaluated and not recommended for consensus standards.

Remaining Criteria: Extent to which the NQF evaluation criteria are met: **H**=high; **M**=moderate; **L**=low. The Committee votes or reaches consensus on ratings.

Recommendation: The Committee/TAP votes on the overall recommendation for endorsement: **Y**=yes, **N**=no, or **A**=abstain.

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
HOE-015-08 Postoperative Respiratory Failure (PSI #11) (Agency for Healthcare Research and Quality)	<p>SC Measure Evaluation criteria: I: Y-15;N-1;A-0 Rationale for ratings (I)/recommendation: I: Although there was some question of the source of estimates for variability 2.3-29.2% and whether wide confidence intervals would negate much variability, because this measure is being used, the committee thought it warranted further evaluation.</p>
	<p>TAP Measure Evaluation criteria: I: Yes (SC) S: H-3;M-6;L-;A- U: H-9;M-;L-;A- F: H-9;M-;L-;A- Recommend for Endorsement w/ Condition: Y-7;N-2;A-0 Rationale for ratings (I, SA, U, F)/recommendation: S: Criterion validity is high, suggesting that the measure is identifying a high number of true positives or true events. The risk-adjustment methodology appears sound, has been used in numerous indicators and settings, and takes into account clustering within hospitals. The indicator is used specifically to examine the quality of care within a specific hospitalization, so that measurement is relatively precise. The measure has been used in several settings with comparable results and high positive predictive validity. Someone questioned whether the false negative rates had been evaluated; however others pointed out that has not been a requirement for testing and this measure has had other appropriate reliability and validity testing. F: Use of administrative data makes feasible. The TAP recommended this measure on the condition that the results of current validation testing are reported as soon as possible. A suggestion also was made that at the time of maintenance review, an assessment of the use the POA indicator be included. Measure Steward Response:</p>
	<p>SC Measure Evaluation criteria: I: Y-15;N-1;A-0 S: H-10;M-6;L-0;A- U: H-11 ;M-5;L-0;A- F: H-12;M-4;L-0;A- Recommend for Endorsement: Y-16;N-0;A-0 Rationale for ratings (I, SA, U, F)/recommendation: The SC agreed that the measure met scientific acceptability, usability and feasibility criteria. S: Because all measures are reviewed under maintenance on a 3-year cycle, it recommended for endorsement with any updated information provided at the time of maintenance review. In response to a question, the developer stated the validation study will</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
	<p>address false negatives, but it is a sampling challenge to do efficiently because relatively infrequent. Also clarified that the variability data should have been 2.3-29.2 per 1000 (not percent).</p> <p>Measure Steward Response:</p>
<p>HOE-008-08 Hospital specific risk-adjusted measure of mortality or one or more major complications within 30 days of a lower extremity bypass (LEB). (Centers for Medicare and Medicaid Services)</p>	<p>SC Measure Evaluation criteria: : I: Y-14;N-1;A-0 Rationale for ratings (I)/recommendation: : I: All sub-criteria were met. One committee member questioned whether it was high enough volume to be considered high impact or best for internal QI only. Another committee member thought it is also an indicator of appropriate pre-operative patient selection. In response to a question regarding variability Bruce Hall, NSQIP stated that 17% experience an event and variability of risk-adjusted predicted:expected ratio is 0.75 to 1.25.</p> <p>TAP Measure Evaluation criteria: I: Yes (SC) S: H-4;M-5;L-;A- U: H-0;M-6;L-3;A- F: H-0;M-4;L-5;A- Recommend for Time-Limited Endorsement: Y-8;N-1;A-0 Rationale for ratings (I, SA, U, F)/recommendation: S: Submission indicates not fully developed and tested and will be completed within 24 months, however development testing reported is quite extensive. Data fields are well defined, but the developer indicated reliability testing would be completed prior to implementation. The TAP questioned whether reliability would hold up when implemented outside of NSQIP's training and auditing and also noted that the risk model would need to be recalibrated. The measure steward noted that NSQIP currently captures about 90% of cases, so would expect relatively few changes. It was clarified that although the measure was developed using NSQIP database, participation in NSQIP is not a requirement for implementation. The measure has a multiple endpoints because of the low occurrence of each event individually, but is not submitted as a composite measure. The reliability of the functional status risk variable was questioned, as well as the validity of RVU as a risk factor. Others commented that accuracy of risk factors overall are less a problem than accuracy of the outcome data. Creatinine>1.2 is a risk factor - should also consider code for dialysis. Developed using 3 years of data, but anticipate computing yearly rates when implemented. The presentation of interval estimates is a strength of the proposed methodology. U: In response to a question whether rates could be improved, the developer stated that they have seen improvement in NSQIP. F: Uses clinical data that until electronic records are available must be collected and reported (now to NSQIP registry, possibly some other mechanism). Participation in NSQIP is not a requirement. Feasibility cannot be entirely evaluated because a national data collection strategy has not yet been proposed. The TAP recommended time-limited endorsement due to development using registry data vs. implementation intended nationally, no report about reliability testing, and need to recalibrate the risk model when implemented nationally. Measure Steward Response: Response regarding RVU: Years ago in the NSQIP the program attempted to control for "procedural complexity" by creating an in-house scale of complexity developed by a panel of experts, but it became apparent that this same information was largely captured already by the CMS designation of work RVUs, with the added advantage that this was an independent body doing the assessments, and that the assessments were updated periodically. It was demonstrated within NSQIP then that the correlation between work RVUs and the in-house "complexity score" was high, and so the complexity score was dropped and the work RVUs were adopted. Again, the aim was to provide some control for procedural complexity, within or across procedure types. Thus, there is now many years of experience using work RVU as a risk adjuster within the NSQIP, and that experience was carried forward into this project. In this vascular project, wRVU continued to demonstrate explanatory value as a risk adjuster (as reflected in the submitted materials). Keep in mind, however, that this LEB measure only deals with a well-defined subset of vascular procedures; controlling for complexity of procedures is always less important within a small procedure subset than it would be for comparisons across large sets of disparate procedures. Nonetheless, the inclusion of wRVU in this vascular measure did contribute to explanatory power.</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
	<p>SC Measure Evaluation criteria: <u>I</u>: Y-14;N-1;A-0 <u>S</u>: H-15; M-1;L-0;A-1 <u>U</u>: H-10;M-6;L-1;A-1 <u>E</u>: H-5;M-9;L-3;A-1</p> <p>Recommend for Time-Limited Endorsement: Y-15;N-1;A-1</p> <p>Rationale for ratings (I, SA, U, F)/recommendation</p> <p>The SC thought the measure was well developed, and concurred with much of the TAP evaluation.</p> <p>S: In response to a question about trauma patients, the steward noted trauma patients are not intended to be included and are not in the NSQIP database. A SC member questioned the reliability of functional status data and another questioned whether other risk factors might enter the model (MI past 6 months) if possibly correlated factors (SGOT, albumin) were removed. The steward responded that functional status has been used in NSQIP and is rigorously defined.</p> <p>U: In response to a comment about understandability of hierarchical modeling, several committee members agreed that understanding a methodology is not essential, as long as the data is presented in a understandable and useful way.</p> <p>F: The steward noted that participation in NSQIP is not a requirement of using this measure. The SC thought that the TAP suggestion for time-limited endorsement probably relates to feasibility of national implementation. The measure requires data after hospitalization that must be obtained from ambulatory records or patient follow-up.</p> <p>Measure Steward Response: the ACS NSQIP has always excluded trauma cases, therefore this entire project has been based on that format. More specifically, any case that activates a trauma resuscitation or work-up is excluded.</p> <p>In regards to data definitions and collection, the data elements were previously specified in our submitted materials as per the standard ACS NSQIP data definitions. Please find the current definitions document as well as a copy of the data collection form attached. The latter is provided to, but not required to be utilized by sites.</p> <p>Your final question addressed item 25 on the submission and reliability testing. Unfortunately, it is not entirely clear to us what this item actually means. If it is in reference to the statistical calculation referred to as reliability, then this information was submitted with our prior materials, and would also be calculated over time during any implementation. If this is in reference to some other measure of "pragmatic" reliability of the measure during implementation, then we are uncertain regarding what measure is being specifically referenced. In either case, our understanding is that part of the reason this measure is being considered for "time limited" approval is so that additional performance information pertaining to the measure could be submitted for review after the measure was implemented for a period.</p>
<p>HOE-009-08 30-day all-cause risk-standardized percutaneous coronary intervention (PCI) mortality rate for patients without ST segment elevation myocardial infarction (STEMI) and without cardiogenic shock (Centers for Medicare and</p>	<p>SC Measure Evaluation criteria: <u>I</u>: Y-15;N-0;A-0</p> <p>Rationale for ratings (I)/recommendation: I: All sub-criteria met. Issues to address in further evaluation: be cognizant of other reporting initiatives and harmonization; whether shock should be separate.</p> <p>TAP Measure Evaluation criteria: <u>I</u>: Yes (SC) <u>S</u>: H-9;M-;L-;A- <u>U</u>: H-6;M-3;L-;A- <u>E</u>: H-3;M-6;L-;A-</p> <p>Recommend for Time-Limited Endorsement: Y-9;N-0;A-0</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: Measures HOE-009 and HOE-010 are basically the same except for the denominator populations (with or without STEMI/cardiogenic shock), which are clearly distinct, both from a clinical standpoint as well as from a data collection standpoint. The measures were discussed and voted on together. The following comments pertain to both HOE-009 and HOE-010.</p> <p>S: Submission indicates not fully developed and tested and will be completed within 24 months, however development testing was reported. Data fields are well defined, but the developer indicated reliability testing would be completed prior to implementation. It was clarified that the measure was developed using NCDR CathPCI registry database, but participation in the registry is not a requirement for implementation. The measure submitted requires matching registry data to Medicare claims and enrollment data. The developer indicated that availability of patient identifier would improve the measure through ability to link with actual outcome (rather than probabilistic matching to outcomes). The TAP agreed that probabilistic matching to endpoint would not be acceptable for</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
Medicaid Services)	<p>a publicly reported measure. The TAP also agreed that 30-day mortality was preferable to in-hospital mortality and that the use of clinical data as in the registry is preferred to administrative data alone. The definition of cardiogenic shock needs reliability testing and may need refinement.</p> <p>U: The endpoint for this measure is easily comprehensible to both the general public as well as to clinicians; it is useful for both consumers and providers.</p> <p>F: Measure is based on an existing registry in which the majority of hospitals that perform PCI already participate and for whom feasibility is high. Those not already participating will need to allocate staff for data collection. Although all data elements are in the electronic registry, they are not currently extracted from an electronic medical record. Feasibility cannot be entirely evaluated because a national data collection strategy has not yet been proposed.</p> <p>The TAP recommended time-limited endorsement due to development using registry data vs. implementation intended nationally, probabilistic matching for testing vs. unique identifiers for implementation, need to recalibrate the risk model when implemented nationally, and no report about reliability testing for key cohort identification and risk adjustment variables.</p> <p>Measure Steward Response: Response regarding additional testing: In our NQF applications for the PCI mortality models, we indicated we would conduct additional testing within 24 months (Page 1, Question D) because CMS plans to refit the models once a national dataset with patient identifiers is assembled for public reporting (this was probably a conservative interpretation of the question). As you know, because we were not able to use direct patient identifiers during the process of measure development, we used a probabilistic match to merge CathPCI registry data with administrative data available on the subset of Medicare patients. The characteristics of patients who matched are virtually identical to those of patients excluded from measure development because they were not matched (Table 5 of the technical report). Accordingly, we are confident that the patients in our analysis are representative of the larger cohort of Medicare patients. In the course of measure implementation, we will have to refit the models using direct identifiers in all PCI patients. However, we would consider these steps part of measure maintenance as opposed to measure development. We have reviewed the criteria for “adequate field testing” set forth in NQF’s guidance on time-limited endorsement and believe that the PCI measures meet these criteria. The measures were developed in a large, representative cohort of PCI patients. Specifically, we analyzed data from more than 125,000 patients undergoing PCI at more than 600 hospitals that submit data to the American College.</p>
	<p>SC Measure Evaluation criteria: <u>I</u>: Y-15;N-0;A-0 <u>S</u>: H-15;M-2;L-0;A- <u>U</u>: H-16;M-0;L-1;A- <u>F</u>: H-0;M-17;L-0;A-</p> <p>Recommend for Endorsement: Y-16;N-0;A-1</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: S: The SC concurred with much of the TAP's evaluation. The two main points of discussion were about 1) including non-STEMI with the non-MI patients or with the STEMI and 2) detecting true cardiogenic shock. The stewards responded that the non-STEMI mortality rate was more similar to non-MI than STEMI and the risk model will adjust for different levels of severity; another consideration was that adding a third category would introduce problems with case volume size. It was discussed that the National Cardiovascular Registry has defined cardiogenic shock consistent with the literature, but in MA they have reviews of designation of cardiogenic shock. The steward suggested that the data could be monitored by identifying hospitals that seem to have an unusual distribution of patients with cardiogenic shock. In regards to the issue raised by SCAI regarding outpatient procedures, the steward noted that inclusion in the measures was not dependent on a hospital admission, just if the PCI was done.</p> <p>The steward clarified that on the submission the future reliability testing was in reference to statistical reliability, not data reliability which has been established. Therefore, the SC recommended for endorsement rather than time-limited endorsement.</p> <p>Note: Because there are competing endorsed and candidate measures on the same topic (NQF# 0133, PCI mortality risk-adjusted; HOE-013, Leapfrog survival predictor for PCI), the committee’s recommendations are conditional on further evaluation that HOE-</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
	<p>009/010 are superior or provide distinctive or additive value to the existing endorsed and candidate measures.</p> <p>Measure Steward Response: The intent has always been to use the proposed PCI measures to publicly report hospital specific 30-day mortality rates of all patients undergoing PCI. Accordingly, the measure will include patients irrespective of admission status and will include outpatient and observation stay patients who have undergone PCI but not been admitted. Patients will be included in the measure denominators based on their undergoing a PCI procedure recorded in the CathPCI or comparable registry, and the 30-day period of assessment will begin immediately thereafter. This is feasible because hospitals will submit data on PCI procedures irrespective of whether the patients is admitted or simply observed following the procedure. Vital status will be determined using an external database such as the Social Security Death Index.</p> <p>As you know, due to the absence of direct patient identifiers in the NCDR CathPCI registry, the PCI mortality measures were developed in a subset of Medicare patients who were admitted following PCI. The models that we developed for use in these measures will be recalibrated (variable coefficients re-estimated) in the larger population of patients. We expect the model to perform equally well in this broader population and therefore do not believe the intent to expand the cohort beyond the inpatient Medicare population necessitates a time limited endorsement.</p>
<p>HOE-010-08 30-day all-cause risk-standardized Percutaneous Coronary Intervention (PCI) mortality rate for patients with ST segment elevation myocardial infarction (STEMI) or cardiogenic shock (Centers for Medicare and Medicaid Services)</p>	<p>SC Measure Evaluation criteria: I: Y-15;N-0;A-0</p> <p>Rationale for ratings (I)/recommendation: See comments for HOE-009.</p> <p>TAP Measure Evaluation criteria: I: Yes (SC) S: H-9;M-;L-;A- U: H-6;M-3;L-;A- F: H-3;M-6;L-;A-</p> <p>Recommend for Time-Limited Endorsement: Y-9;N-0;A-0</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: See comments for HOE-009</p> <p>Measure Steward Response: See comments for HOE-009.</p> <hr/> <p>SC Measure Evaluation criteria: I: Y-15;N-0;A-0 S: H-15;M-2;L-;A- U: H-16;M-0;L-1;A- F: H-0;M-17;L-0;A-</p> <p>Recommend for Endorsement: Y-16;N-0;A-2</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: See comments for HOE-009.</p> <p>Measure Steward Response: See comments for HOE-009.</p>
<p>HOE-013-08 Replaced by HOE-019 to 024 Survival Predictor (6 individual mortality measures – CABG, AVR, PCI, AAA, Esophagectomy, Pancreatectomy) (Leapfrog Group)</p>	<p>SC Measure Evaluation criteria: I: Y-14;N-0;A-1</p> <p>Rationale for ratings (I)/recommendation: I: All components are NQF-endorsed so already determined to be important. Issues to address in further evaluation: weighting of components, how to sort out elective from emergent procedures (e.g., AAA). One committee member questioned inclusion of volume component in an outcome measure; and another noted the volume-outcome relationship for some conditions/procedures is controversial. It was noted that NQF has a Composite Steering Committee that has been developing evaluation criteria that will be provided to the committee and TAP.</p> <p>TAP Measure Evaluation criteria: I: Yes (SC) S: H-0;M-3;L-6;A- U: H-1;M-5;L-3;A- F: H-0;M-0;L-8;CA-1;A-</p> <p>Recommend for Endorsement: Y-0;N-8;A-1</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: Although only one measure submission form was submitted that referred to a survival predictor and listed 12 component measures, the measure steward clarified that there are 6 separate mortality measures (CABG, AVR, PCI, AAA, Esophagectomy, Pancreatectomy).</p> <p>S: Although the TAP agreed that the Bayesian methodology and modeling is elegant and cutting edge, it agreed it was not ready for</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
	<p>endorsement. The proposed composite measures are a weighted average of a facility's mortality and the expected mortality given its volume ($E[m]=a+b*\log[\text{volume}]$). Facilities with a small number of cases then get weighted more heavily towards the expected mortality given their volume. This expected mortality is likely to be higher than the mean across all facilities since lower volumes are generally associated with higher mortality. The TAP noted the controversy surrounding the volume-outcome relationship and questioned the premise of using a volume-predicted mortality rate as a component of the composite. Although the methodology employed by this measure was recently published in a prestigious journal (Medical Care), the panel noted that publication of a single article often marks the beginning, not the end, of a discussion of a controversial subject. The panel expects that this paper will trigger much discussion as well as the publication of counter-examples and critiques, and that this process will take some time before consensus is reached on the volume-outcome relationship. Another issue identified was the lack of standardization regarding risk adjustment - the specifications allow for either risk-adjusted or raw mortality rates. The developer stated risk adjustment makes no difference in predicting future risk-adjusted rate. The competing NQF-endorsed mortality measures are all risk-adjusted.</p> <p>U: Because there are already NQF-endorsed mortality measures for the six procedures, the question is whether these represent additive value or superior methodology. The measure steward noted that the current NQF-endorsed cardiovascular measures from STS and ACC/AHA are not currently publicly reported.</p> <p>The TAP did not think these measures were ready to replace the existing endorsed measures.</p> <p>Measure Steward Response: The steward stated it will use risk-adjusted rates for hospitals that have them available and the survival predictor with unadjusted rates for those that don't. The steward noted that the survival predictor did predict future risk-adjusted mortality, possibly because only including elective procedures.</p>
	<p>SC Measure Evaluation criteria: I: Y-14;N-0;A-1 S: H-1;M-12;L-2;A- U: H-1;M-8;L-6;A-1 F: H-8;M-7;L-0;A-</p> <p>Recommend for Endorsement: Y-9;N-6;A-0</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: The SC was divided on recommending these measures. The primary reasons for recommending these measures despite the TAP's reservations were 1) even though there is controversy in general regarding volume-outcome relationships, the steward noted that the relationship is established for these procedures and 2) they allow reporting on hospitals with small case volumes (the steward indicated that 1 case can be reported). The minority position against these measures was based on 1) agreement with the TAP on the controversy regarding volume-outcome relationships including the role of surgeon vs. hospital; and therefore should not be the premise for the rate for small volume providers, which will be characterized primarily by the volume-predicted rate and 2) the measures only provide a slight marginal benefit over the existing measures by being able to report on small volume providers and whether reporting a rate for a provider with as few as 1 case, which would be based primarily on a volume-predicted rate, is useful information. One member also commented that by removing the emergent cases, you are perhaps missing an assessment of the skill of those providers.</p> <p>U: The specifications for the 3 endorsed AHRQ QI measures was handed out at the meeting, but the detailed specifications for the other 3 endorsed measures had not been obtained as of the meeting time, and the SC was not able to fully compare the measures. Committee members thought that being able to measure small providers was an advantage.</p> <p>Note: Because there are competing endorsed measures on the same topic, the committee's recommendation is conditional on further evaluation that the candidate measures are superior to the existing endorsed measures or provide distinctive or additive value.</p> <p>Existing measures: NQF# 0359, Abdominal Aortic Artery (AAA) Repair Mortality Rate (IQI 11) (risk adjusted); NQF# 0360, Esophageal Resection Mortality Rate (IQI 8) (risk adjusted); NQF# 0365, Pancreatic Resection Mortality Rate (IQI 9) (risk adjusted); NQF# 0133, PCI mortality risk-adjusted; NQF# 0119, Risk-Adjusted Operative Mortality for CABG; NQF# 0120, Risk-Adjusted Operative Mortality for Aortic Valve Replacement. Candidate measures: HOE-009/010, 30-day All-cause Risk-standardized PCI</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
	<p>Mortality. Measure Steward Response: Submitted each measure individually.</p>
<p>HOE-004-08 RISK-ADJUSTED 30-DAY READMISSION RATE FOR HEART FAILURE (Health Benchmarks, Inc)</p>	<p>SC Measure Evaluation criteria: I: Y-15;N-0;A-0 Rationale for ratings (I)/recommendation: I: SC already agreed on importance of readmission in Phase I. Issues to address in further evaluation: attention to left ventricular assistive devices and transplant; comparison to other endorsed measures.</p> <p>TAP Measure Evaluation criteria: I: Yes (SC) S: H-0;M-8;L1-;A- U: H-0;M-7;L-2;A- F: H-0;M-8;L-1;A- Recommend for Endorsement: Y-0;N-9;A-0 Rationale for ratings (I, SA, U, F)/recommendation: S: A number of issues were identified. It appears the risk models are fit to each plan rather than one that applies to all hospitals. An exclusion (hospice before & up to 30 days after) and risk factor (discharge to nursing home 1-30 days after) occur after discharge and may inappropriately exclude or adjust for outcomes that are the result of care. There is conflicting information on how age is used (dichotomous, categorical). The comorbidity index includes some of the other individual risk factors (e.g., COPD, renal failure). U: Although this measure would apply to potentially all patients vs. the competing endorsed measure that applies only to Medicare patients, it would be limited to health plans because of the need to link claims over time. The TAP agreed this measure was not strong scientifically. Measure Steward Response: Issue 1: The risk model should be fitted to all plans rather than to each plan. Response: We fit the risk model to all plans and present the results and statistics regarding the model in table 1 and 2. Issue 2: Discharged to nursing home should not be included as a variable for risk-adjustment. Response: This measure is designed to use administrative claims data only to maximize ease of use and widest adoption. Administrative claims data do not capture direct information regarding severity of heart failure; thus we used discharged to nursing home as proxy measure for more severe heart failure. Issue 3: Members with heart failure and on hospice before and after discharge may be inappropriately excluded. Response: We conceptually excluded patients who receive hospice on discharge because members on hospice most likely have end stage heart failure and would be admitted only for palliative treatment. Less than 1% of our commercially insured sample received hospice during the time period specified and including and excluding these patients made no difference in the results. Issue 4: There is conflicting information on how age is used. Response: Thank you for this feedback. We revisited the age variable in our model and determined that the best way to use age is as a continuous variable with a quadratic term to capture the nonlinear relationship (Table 1). Issue 5: Comorbidity index includes some of the other individual factors (e.g., COPD and renal failure). Response: We apologize for this misunderstanding. We used the modified Elixhauser Comorbidity Index in which we excluded CHF, renal failure, and COPD calculation of the index.</p> <p>SC Measure Evaluation criteria: I: Y-15;N-0;A-0 S: H-8;M-8;L-0;A- U: H-5;M-7;L-3;A-1 F: H-7;M-5;L-3;A-1 Recommend for Endorsement: Y-12;N-2;A-2 Rationale for ratings (I, SA, U, F)/recommendation: S: In response to a question regarding whether there was adequate case volumes of CHF patients <65 who are hospitalized, a SC member stated the average age for CHF in a 3,000-patient clinical trial is 59. The steward also stated it had run its model on a dataset with Medicare Advantage patients. The committee thought a measure for all payers and all ages was appropriate. The primary reason for approving this measure was that it includes all payers and ages and the issues raised by the TAP were addressed. It was noted that the steward should not refer to the Elixhauser index if it has been modified. Note: Because there is a competing endorsed measure on the same topic (NQF# 0330, 30-Day All-Cause Risk Standardized</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
	<p>Readmission Rate Following Heart Failure Hospitalization risk adjusted), the committee’s recommendation is conditional on further evaluation that the candidate measure is superior or provides distinctive or additive value to the existing endorsed measure.</p> <p>Measure Steward Response: Provided updated submission and information regarding age distributions.</p> <p>*Item 26 – Is it correct that you used only patients>65 from the Pharmetrics data? Yes, initially we only used patients>65 from the Pharmetrics data because we wanted to validate our model using a large sample from Medicare population. However, when NQF asked for the age distribution of the members in our data base we submitted the statistics for all the age ranges for the Pharmetrics data. If NQF is interested in the results we can request that this model be ran on all eligible patients in Pharmetrics data.</p> <p>*Item #28 – Please confirm that the # of cases are # of discharges with heart failure. Of these cases of heart failure, what percentage is below 65? Yes, In Item 28, the # of cases is the # of discharges with heart failure. Of these cases of heart failure, 17% are below 65. This is different from the overall percentage estimate from HBI data because Item 28 only used data from one health plan.</p> <p>*You gave me the age distribution for your datasets (attached #1) – are these distributions for the entire data set or for CHF admissions in your dataset? The age distribution of the HBI data set presented in Table 1 is the distribution for the entire HBI CHF admissions in 2006. Although we validated our model on patients > 65 in the Pharmetrics data set, in table 2 we present the entire distribution of patients in the 2007 Pharmetrics data set so that NQF committee members can see the distribution of heart failure patients by age groups and plan type in one of the largest commercial data set in the US.</p> <p>*The measure specifications specifically refer to “members” so is it correct to say the measure is designed for use by health plans? In other words, would the measure computed for each hospital only include patients of a particular health plan (vs. all patients at the hospital, vs. Medicare FFS patients)? The easiest way to apply this measure is to data from a specific health plan. The measure needs one year of pre-discharge data to calculate variable for risk adjustment. To include the patient in the denominator, we must have evidence that we have complete administrative claims data for this patient in the 365 days prior to discharge and 30 days after discharge. This measure will also work If we have information from multiple payers (multiple health plans) in the same geographic area which serves the hospitals. However, some additional work needs to be done to align the hospital identifications among the different health plans.</p>
<p>HOE-011-08 Measure of the Occurrence of deep-vein thrombosis/pulmonary embolism (DVT/PE) Following Hip or Knee Replacement Surgery (Johnson & Johnson Health</p>	<p>SC Measure Evaluation criteria: I: Y-14;N-1;A-1 Rationale for ratings (I)/recommendation: I: DVT is an important topic of measurement and relates to NPP goal. Information was provided on impact, but not variability in performance. Issues to address in further evaluation: it is untested and there are already many measures related to DVT; 30-day time frame</p> <p>TAP Measure Evaluation criteria: I: Yes (SC) S: H-0;M-0;L9-;A- U: cannot determine F: cannot determine Recommend for Time-Limited Endorsement: Y-0;N-9;A-0 Rationale for ratings (I, SA, U, F)/recommendation: S: Measure is untested. The measure is intended to identify treatment for DVT/PE 30 days after discharge; however the specifications do not provide any detail for ambulatory coding and linking index hospitalization to post-discharge hospital and ambulatory claims. No risk adjustment strategy is planned because the steward states that DVT/PE is considered "preventable for all patient risk profiles"; however, in item #19 of the submission form, the steward identified factors associated with disparate outcomes including cancer, obesity, age, previous VTE, oral contraception, which indicates</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
Care Systems, Inc.)	<p>the need for some method of risk adjustment or risk stratification. F: Only feasible if able to link claims across time and settings. Measure Steward Response: SC Measure Evaluation criteria: <u>I</u>: Y-14;N-1;A-1 <u>S</u>: H-0;M-0;L-15;A-1 <u>U</u>: H-0;M-5;L-9;A-1 <u>F</u>: H-2; M-5;L-7;A-1 Recommend for Time-Limited Endorsement: Y-0;N-14;A-1 Rationale for ratings (I, SA, U, F)/recommendation: S: The SC agreed with the TAP that the measure specifications need to be more precise in order to implement and that a risk adjustment strategy is needed. A SC member discussed that the majority of events following hospitalization are clinically silent and this measure only looks at those that are identified, which is a smaller percentage. Others noted that practice guidelines are discordant. One member questioned whether this was an appropriate topic for the 30-day time window; another suggested that would encourage better surveillance and prophylaxis. F: Health plans and CMS do have the ability to link claims across time and settings, so they would be able to implement such a measure. The SC agreed this measure as specified was not ready for endorsement, but expressed they would like to see this measure brought back to NQF in the future after testing. Measure Steward Response:</p>
HOE-014-08 Postoperative Hemorrhage and Hematoma (PSI #9) (Agency for Healthcare Research and Quality)	<p>SC Measure Evaluation criteria: <u>I</u>: Y-5;N-11;A-0 Rationale for ratings (I)/recommendation: I: Although relates to NPP safety goal, it is infrequent (96 deaths); there is little variability (2.3 to 2.9) and large number with 0 rate. Although it is an outcome measure, it would have little improvement impact.</p>
HOE-005-08 Postprocedural Stroke or Death in Asymptomatic Patients undergoing Carotid Angioplasty and Stenting (Northwestern University, Society for Vascular Surgery)	<p>SC Measure Evaluation criteria: <u>I</u>: Y-4;N-11;A-0 Rationale for ratings (I)/recommendation: I: Although preventing stroke and death are worthwhile goals, there are no data on this relatively new procedure and the procedure itself is not yet supported outside of clinical trials.</p>
HOE-017-08 Postoperative Stroke or Death in Asymptomatic Patients undergoing	<p>SC Measure Evaluation criteria: <u>I</u>: Y-11;N-1;A-3 Rationale for ratings (I)/recommendation: TAP Measure Evaluation criteria: <u>I</u>: Y-9;N-0;A- <u>S</u>: H-;M-5;L-4;A- <u>U</u>: H-0;M-2;L-7;A- <u>F</u>: H-0;M-0;L-9;A- Recommend for Time-Limited Endorsement: Y-0;N-9;A-0 Rationale for ratings (I, SA, U, F)/recommendation: I: This measure submission had been inadvertently missed. The TAP agreed it</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
Carotid Endarterectomy (Society for Vascular Surgery)	<p>met the importance criterion. These are important outcomes and measure also would encourage selection of appropriate patients for the procedure.</p> <p>S: The measure is untested. The measure would require physician claims be linked to hospital claims in order to have the information in the G-code that indicates the patient was asymptomatic for a year prior to the procedure. Although the measure would not need risk adjustment if restricted to the asymptomatic patients, testing of the reliability and validity of the G-code, especially for under-reporting is necessary. The TAP also did not think that a cumulative lifetime rate for individual physicians was a sound approach for performance measurement and that other approaches to deal with small volume should be explored (e.g., rolling time periods).</p> <p>F: G-code not yet established and G-codes not used in hospital claims.</p> <p>These issues do not warrant granting time-limited endorsement – the measure should be tested and then brought back to NQF.</p> <p>Measure Steward Response:</p> <p>SC Measure Evaluation criteria: I: Y-11;N-1;A-3 S: H-0;M-0;L-15;A- U: H-0;M-0;L-15;A- F: H-0;M-0;L-15;A- Recommend for Time-Limited Endorsement: Y-0;N-15;A-0</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: The SC agreed with the issues identified by the TAP.</p> <p>S: The SC discussed whether there should be some risk adjustment even among asymptomatic patients. The steward submission indicates that the measure is targeted for asymptomatic patients because practice guidelines recommend CEA only be performed in asymptomatic patients and it is incumbent on the surgeon to only select patients for this prophylactic and elective operation who will have a low stroke or death rate.” Committee members identified that factors such as age and other comorbidities such as renal failure, diabetes, etc. affect risk.</p> <p>The SC expressed they would like to see this measure brought back to NQF in the future after refinement and testing.</p> <p>Measure Steward Response:</p>
HOE-016-08 RISK-ADJUSTED COMPLICATION LIKELIHOOD FOR SURGERIES: APPENDECTOMY AND CHOLECYSTECTOMY (Health Benchmarks, Inc)	<p>SC Measure Evaluation criteria: I: Y-5;N-9;A-1</p> <p>Rationale for ratings (I)/recommendation: I: This is a fairly common procedure and relates to NPP safety goal; however, the quality problem was not demonstrated. In response to question of why these procedures were chosen, the measure steward indicated it was because they are relatively common.</p>
HOE-018-08 Inpatient Co-morbidity Adjusted Complication Index (Premier, Inc)	<p>SC Measure Evaluation criteria: I: Y-18;N-0;A-0</p> <p>Rationale for ratings (I)/recommendation:</p> <p>TAP Measure Evaluation criteria: I: Y-8;N-0;A-1 S: H-0;M-8;L-1;A- U: H-0;M-0;L-9;A- F: H-1;M-5;L-3;A- Recommend for Endorsement: Y-0;N-9;A-0</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: Please note that the measure steward was notified that for this project, specific measures need to be proposed and evaluated. Although the Premier classification system facilitates drilling down into the data for various levels of analyses, the measures being evaluated for potential endorsement are for total complications or total severe complications across all hospitalized patients. Measures HOE-018 and HOE-006 are both measures of complications using the same methodology. HOE-018 includes all complications; HOE-006 includes severe complications. The following comments pertain to both HOE-006 and HOE-018.</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
	<p>I: Measure HOE-018 submission had been inadvertently missed and not previously reviewed by the Steering Committee. The TAP agreed it met the importance criterion.</p> <p>S: A primary concern was the replicability of the classification system. The definition of what constitutes a complication is dependent upon an evaluation of principal-secondary diagnosis pairs selected by volume and reviewed by physician panels using modified Delphi consensus techniques to determine the probability that the secondary dx is a complication rather than a comorbidity. Complications also are classified by severity on a 5-point Likert scale (A-E) by internal panels of clinicians and those rated D&E are used to denote the severe complications for HOE-006. The risk adjustment model includes race and income variables, which the NQF evaluation criteria suggest should not be used in risk adjustment. The developer stated these are considered proxies for access to care. The risk model also includes valid procedure that occurs during the hospitalization ("certain procedures can serve as effective proxies for lab reports and treatment history that are not available in the current database, as well as for other unobservable critical factors.") and discharge status (which occurs after care is provided). Risk model performance metrics for a development and validation sample were not provided. The number of diagnoses included in a data source (e.g., CA-25, MEDPAR-9) can affect rates of complications unless hospitals only compared within the same data source. Only face validity is addressed, and there was no testing to determine if variability in scores reflects variability in complication rates or in coding practices. The TAP discussed that use of administrative ICD-9 codes to identify an outcome such as all complications vs. variables used in risk models, necessitates an understanding of the reliability of those data.</p> <p>U: The TAP discussed whether a global complications measure based on diagnosis codes can be used for public reporting because of the reliability and validity issues identified above, although using such methods as screening tools for quality improvement activities might be helpful. In addition, the overall scores subsume many different complications so that the type of complications could differ greatly from one hospital to another. There also was some discussion of whether a global complications measure can be used for quality improvement without data on the specific complications. However, risk-adjusted complication rates from coded data could identify situations that require further investigation. The classification system used to compute the measures also can be used in the QI investigation to identify patients or various groups of patients (e.g., by diagnosis) with the complications. The developer stated that hospitals would need to do that analysis on their own and it could be done in a simple spreadsheet.</p> <p>F: The measure is based entirely on administrative data. The measure steward plans to make the measure freely available. The TAP agreed that such a measure is useful for screening and the system a useful tool for QI investigations, but it is not ready for publicly benchmarking performance.</p> <p>Measure Steward Response: The measure steward submitted an additional 42 pages of materials and a letter regarding concerns about process.</p>
	<p>SC Measure Evaluation criteria: I: Y-18;N-0;A-0 S: H-0;M-10;L-8;A- U: H-1;M-13;L-4;A- F: H-7;M-11;L-0;A-</p> <p>Recommend for Endorsement: Y-5;N-11;A-2</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: The following comments pertain to both HOE-006 and HOE-018.</p> <p>S: The Delphi process may be an appropriate tool for reaching consensus, however, SC members expressed concern about 1) the lack of reliability and validity testing to verify the probabilities that a secondary diagnosis is a complication, 2) the use of a subjective process in areas where evidence is available, and 3) not using the POA indicator. The steward indicated it was planning to use POA in the future. The committee discussed that the POA is important to distinguish co-morbidities, but is not yet ready for implementation. The committee also noted the use of race and SES as risk factors and the NQF criteria suggesting stratification rather than risk adjustment. The steward indicated it found differences in complication rates by race and SES, but felt that was related to access to care rather than differences in care provided. Note: an analysis by the measure developer submitted in support of the measures states, "the process of care differences that we and others have observed suggest that some of the difference in complication rates can be attributable to</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
	<p>systematic racial differences in the way patients are treated after admission to a hospital." (Kroch, Eugene A., et. al. "Racial Disparities in Inpatient Complication and Mortality Rates." Academy Health Presentation. June 2003)</p> <p>U: Most committee members agreed that the aggregate rate of complications is less useful than more specific measures for both consumers and providers; however, a member noted that such an aggregate score would be of interest to purchasers. The SC agreed that the system is useful for screening and detailed analysis to identify specific problems for quality improvement, but not public reporting.</p> <p>F: The measure is based on administrative claims data and does not require additional data collection. Premier has indicated that if endorsed, it would make use of the measure free of charge and will provide a mechanism for entering data online. If hospitals want detailed reports to disaggregate the data to assist with targeting quality improvement it would need to subscribe to Premier's quality improvement system.</p> <p>Measure Steward Response:</p>
<p>HOE-006-08 Inpatient Co-morbidity Adjusted Morbidity Index (Premier, Inc)</p>	<p>SC Measure Evaluation criteria: <u>I</u>: Y-12;N-4;A-0</p> <p>Rationale for ratings (I)/recommendation: I: Relates to NPP goals and is a relevant outcome for patients with variability in performance. Issues to address in further evaluation: usability due to broad focus; adequacy of risk adjustment; ability to identify co-morbid conditions; lack of specificity for target population – can be applied to any population, but NQF only endorses discrete, specific measures.</p> <p>TAP Measure Evaluation criteria: <u>I</u>: Yes (SC) <u>S</u>: H-0;M-3;L-6;A- <u>U</u>: H-0;M-0;L-9;A- <u>F</u>: H-1;M-5;L-3;A-</p> <p>Recommend for Endorsement: Y-0;N-9;A-0</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: See comments for HOE-018.</p> <p>Measure Steward Response:</p> <p>SC Measure Evaluation criteria: <u>I</u>: Y-12;N-4;A-0 <u>S</u>: H-0;M-13;L-5;A- <u>U</u>: H-1;M-13;L-4;A- <u>F</u>: H-6;M-10;L-1;A-</p> <p>Recommend for Endorsement: Y-7;N-10;A-1</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: See comments for HOE-018.</p> <p>Measure Steward Response:</p>
<p>HOE-007-08 3M™ Potentially Preventable Complications (PPCs) (3M Health Information Systems)</p>	<p>SC Measure Evaluation criteria: <u>I</u>: Y-18;N-0;A-0</p> <p>Rationale for ratings (I)/recommendation: I: The SC agreed with the TAP that this measure met the Importance criterion.</p> <p>TAP Measure Evaluation criteria: <u>I</u>: Y-9;N-0 <u>S</u>: H-0;M-4;L-5;A- <u>U</u>: H-0;M-6;L-3;A- <u>F</u>: H-0;M-2;L-7;A-</p> <p>Recommend for Endorsement: Y-0;N-9;A-0</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: Please note that the measure steward was notified that for this project, specific measures need to be proposed and evaluated. Although the 3M classification system facilitates drilling down into the data for various levels of analyses, the measure being evaluated for potential endorsement is for total complications across all hospitalized patients.</p> <p>I: This measure was not previously reviewed by the SC. The TAP agreed it met the importance criterion.</p> <p>S: This measure builds on the AHRQ Patient Safety Indicators (PSI) and the Complications Screening Program (CSP). The measure will be sensitive to present-on-admission (POA) coding practices, and the developers point out that hospitals have 2 incentives to increase POAs: 1) to decrease complication rate and 2) increase severity of illness. It was developed using CA data where POA has been implemented. The number of diagnoses included in a data source (e.g., CA-25, MEDPAR-9) can affect rates of complications unless hospitals are only compared within the same data source. A TAP member noted that CA data tends to be quite different and that validation with other data sets from other states or a national set would be desirable. Only face validity is addressed, and there was no testing to determine if variability in scores reflects variability in complication rates or in coding practices. Risk adjustment is accomplished by indirect standardization using APR DRGs further subdivided by 4 severity of illness subclasses and 4 risk of</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
	<p>mortality subclasses (developed by and iterative process of formulating clinical hypotheses and then testing the hypotheses with historical data). The TAP discussed that use of administrative ICD-9 codes to identify an outcome such as all complications vs. variables used in risk models, necessitates an understanding of the reliability of those data. It agreed that POA coding will assist with distinguishing complications from co-morbidity; however POA coding is relatively recent in many states and measure scores are subject to variability in coding practices.</p> <p>U: The TAP discussed whether a global complications measure based on diagnosis codes can be used for public reporting because of the issues identified above, although using such methods as screening tools for quality improvement activities might be helpful. In addition, the overall scores subsume many different complications so that the type of complications could differ greatly from one hospital to another. There also was some discussion of whether a global complications measure can be used for quality improvement without data on the specific complications. However, risk-adjusted complication rates from coded data could identify situations that require further investigation. The classification system used to compute the measures also can be used in the QI investigation to identify patients or various groups of patients (e.g., by diagnosis) with the complications.</p> <p>F: The measure is based entirely on administrative data. The measure steward intends to charge for use of the measure, which would require both the PPC system and APR DRGs (stated PPC is roughly half the cost of APR DRG). The TAP agreed that such a measure is useful for screening and the system useful tool for QI investigations, but it is not ready for publicly benchmarking performance.</p> <p>Measure Steward Response: In response to the issue regarding the effect of the number of secondary diagnoses in a data set, the steward indicated that you would need to compute the reference norm for the risk adjustment method and make comparisons by data set. The steward also noted that the historical problems related to POA coding were related to inadequate guidelines and now that CMS is requiring POA coding, there has been a lot of work on laying out precise coding guidelines. 3M submitted a letter with concerns about process.</p> <p>SC Measure Evaluation criteria: <u>I</u>: Y-18;N-0;A-0 <u>S</u>: H-1;M-15;L-2;A- <u>U</u>: H-4;M-13;L-1;A- <u>F</u>: H-2;M-13;L-3;A- Recommend for Endorsement: Y-8;N-10;A-0</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: S: Although this measure uses the POA indicator, which helps distinguish a complication from a pre-existing co-morbidity, the SC did not think POA is ready for use in a measure suitable for public reporting. It also discussed that it expects coding to improve since it is now required on the UB 04 claim form and the coding guidelines have been improved. It was noted that although this classification system did not use a Delphi process as in the Premier measures, it is based on clinical panels deciding what was a potentially preventable complication. SC members expressed concern about the lack of reliability and validity testing to verify the that complications identified are in fact complications.</p> <p>U: Most committee members agreed that the aggregate rate of complications is less useful than more specific measures for both consumers and providers; however, a member noted that such an aggregate score would be of interest to purchasers. The SC agreed that the system is useful for screening and detailed analysis to identify specific problems for quality improvement, but did not reach consensus on its usefulness for public reporting.</p> <p>F: The measure is based on administrative claims data and does not require additional data collection. 3M will charge for the use of its measure, but that also includes the entire system, information and support.</p>
<p>HOE-012-08 3M™ Potentially Preventable Readmissions</p>	<p>SC Measure Evaluation criteria: <u>I</u>: Yes Rationale for ratings (I)/recommendation: I: SC already agreed on importance of readmission in Phase I.</p> <p>TAP Measure Evaluation criteria: <u>I</u>: Yes (SC) <u>S</u>: H-0;M-2;L-7;A- <u>U</u>: H-0;M-0;L-9;A- <u>F</u>: H-0;M-1;L-8;A- Recommend for Time-Limited Endorsement: Y-0;N-9;A-0</p>

Meas# / Title/ (Owner)	TAP/Steering Committee Discussion/Evaluation
(PPRs) (3M Health Information Systems)	<p>Rationale for ratings (I, SA, U, F)/recommendation: Please note that the measure steward was notified that for this project, specific measures need to be proposed and evaluated. Although the 3M classification system facilitates drilling down into the data for various levels of analyses, the measure being evaluated is for total preventable readmissions across all hospitalized patients.</p> <p>S: Although there is some appeal to isolating preventable readmissions, the TAP questioned the reproducibility and validity of the designation of preventable readmissions by clinical panels (the developer indicated “each of the 98,596 cells contain a specification of whether the combination of the base APR DRG for the Initial Admission and for the readmission were clinically-related and therefore potentially preventable”). A question also was raised about the stability of the empiric estimates for the PPRs based on one state (FL). The submission form indicates any risk adjustment method could be used but APR DRGs is recommended; however, one method would need to be specified to result in a standard measure. A limitation of the risk adjustment method is reliance on ICD-9 codes without POA indicator and whether can adequately distinguish what was present at the start of care from conditions that developed during care.</p> <p>U: There is an NQF-endorsed risk-adjusted measure for all readmissions. Comparison of results and rankings from this candidate measure of preventable readmissions with the endorsed risk-adjusted all readmission measure is needed to justify the complexity of this measure.</p> <p>F: The measure is based entirely on administrative data. The measure steward intends to charge for use of the measure, which would require both the PPR system and APR DRGs (stated PPR is roughly half the cost of APR DRG).</p> <p>Measure Steward Response: 3M submitted a letter with concerns about process.</p>
	<p>SC Measure Evaluation criteria: <u>I</u>: Yes-see 004, Phase I <u>S</u>: H-2;M-11;L-4;A- <u>U</u>: H-2;M-12;L-3;A- <u>F</u>: H-4;M-8;L-5;A-</p> <p>Recommend for Endorsement: Y-7;N-11;A-0</p> <p>Rationale for ratings (I, SA, U, F)/recommendation: S: The SC agreed that risk adjustment, time window, and readmission to same hospital or any hospital need to be standardized in the measure specifications and should not be up to any implementer - the steward seems willing to do that. In regards to the 98,000-cell matrix, the steward noted that 2/3 were assessed as not causally related. The SC also expressed concern with reliability of the judgment process used to determine if a readmission was related to the prior admission.</p> <p>U: The steward stated that FL is now publicly reporting on a PPR-based set of reports where in the past it used an all-cause readmission that the providers objected to and did not improve. The SC noted the utility of the "3M system" for quality improvement, but expressed concern that the specific aggregate measure submitted for consideration would not be that useful for consumers or providers; however, one SC member noted it could be useful to purchasers.</p> <p>F: Administrative claims data are feasible to use. However, not all states have a database like FL and even though each payer would have the data for its patients, that would not provide a complete picture for any hospital. The fees for using the measure makes the feasibility score somewhat lower, but some members did not think it was a big issue.</p>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.0

August 2008

The measure information you submit will be shared with NQF’s Steering Committees and Technical Advisory Panels to evaluate measures against the NQF criteria of importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Four conditions (as indicated below) must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, importance to measure and report, in order to be evaluated against the remaining criteria. References to the specific measure evaluation criteria are provided in parentheses following the item numbers. Please refer to the *Measure Evaluation Criteria* for more information at www.qualityforum.org under Core Documents. Additional guidance is being developed and when available will be posted on the NQF website.

Use the tab or arrow (↓→) keys to move the cursor to the next field (or back ←↑). There are three types of response fields:

- drop-down menus - select one response;
- check boxes - check as many as apply; and
- text fields - you can copy and paste text into these fields or enter text; these fields are not limited in size, but in most cases, we ask that you summarize the requested information.

Please note that URL hyperlinks do not work in the form; you will need to type them into your web browser.

Be sure to answer all questions. Fields that are left blank will be interpreted as no or none. Information must be provided in this form. Attachments are not allowed except when specifically requested or to provide additional detail or source documents for information that is summarized in this form. If you have important information that is not addressed by the questions, they can be entered into item #48 near the end of the form.

For questions about this form, please contact the NQF Project Director listed in the corresponding call for measures.

CONDITIONS FOR CONSIDERATION BY NQF	
	Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards.
A (A)	Public domain or Intellectual Property Agreement signed: Public domain - IP agreement not required (If no, do not submit) <i>Template for the Intellectual Property Agreement is available at www.qualityforum.org under Core Documents.</i>
B (B)	Measure steward/maintenance: Is there an identified responsible entity and process to maintain and update the measure on a schedule commensurate with clinical innovation, but at least every 3 years? Yes, information provided in contact section (If no, do not submit)
C (C)	Intended use: Does the intended use of the measure include BOTH public reporting AND quality improvement? Yes (If no, do not submit)
D (D)	Fully developed and tested: Is the measure fully developed AND tested? Yes, fully developed and tested (If not tested and no plans for testing within 24 months, do not submit)

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.0

August 2008

	(for NQF staff use) NQF Review #: HOE-015-08 NQF Project: Hospital Outcomes and Efficiency
MEASURE SPECIFICATIONS & DESCRIPTIVE INFORMATION	
1	Information current as of (date- MM/DD/YY): 08/21/2008
2	Title of Measure: Postoperative Respiratory Failure (PSI #11)
3	Brief description of measure ¹ : Number of adult patients with postoperative respiratory failure per eligible elective admissions
4 (2a)	<p>Numerator Statement: Discharges among cases meeting the inclusion and exclusion rules for the denominator with EITHER</p> <p>1) ICD-9-CM codes for acute respiratory failure (518.81, 518.84) in any secondary diagnosis field OR</p> <p>2) ICD-9-CM codes for reintubation procedure as follows:</p> <ul style="list-style-type: none"> • (96.04) one or more days after the major operating room procedure code • (96.70 or 96.71) two or more days after the major operating room procedure code • (96.72) zero or more days after the major operating room procedure code <p>Time Window: In-hospital</p> <p>Numerator Details (Definitions, codes with description): See above</p>
5 (2a)	<p>Denominator Statement: All elective* surgical discharges age 18 and older defined by specific DRGs and an ICD-9-CM code for an operating room procedure.</p> <p>Time Window: In-hospital</p> <p>Denominator Details (Definitions, codes with description): Elective is defined as admission type recorded as elective or scheduled</p>
6 (2a, 2d)	<p>Denominator Exclusions: Exclude cases:</p> <ul style="list-style-type: none"> • with preexisting (principal diagnosis or secondary diagnosis present on admission, if known) acute respiratory failure • with ICD-9-CM diagnosis code of neuromuscular disorder • where a procedure for tracheostomy is the only operating room procedure. • where a procedure for tracheostomy occurs before the first operating room procedure. <p>Note: If day of procedure is not available in the input data file, the rate may be slightly lower than if the information was available.</p> <ul style="list-style-type: none"> • MDC 14 (pregnancy, childbirth, and puerperium) • MDC 4 (diseases/disorders of respiratory system) • MDC 5 (diseases/disorders of circulatory system) <p>Denominator Exclusion Details (Definitions, codes with description):</p> <p>ICD-9-CM Tracheostomy procedure codes:</p> <p>3121 MEDIASTINAL TRACHEOSTOMY</p> <p>3129 OTHER PERM TRACHEOSTOMY</p> <p>3174 REVISION OF TRACHEOSTOMY</p> <p>ICD-9-CM Neuromuscular Disorder codes:</p> <p>3570 AC INFECT POLYNEURITIS</p> <p>35800 MYSTHNA GRVS W/O AC EXAC</p> <p>35801 MYASTHNA GRAVS W AC EXAC</p>

¹ Example of measure description: Percentage of adult patients with diabetes aged 18-75 years receiving one or more A1c test(s) per year.
NQF Measure Submission Form, V3.0

3581 MYASTHENIA IN OTH DIS
 3582 TOXIC MYONEURAL DISORDER
 3588 MYONEURAL DISORDERS NEC
 3589 MYONEURAL DISORDERS NOS
 3590 CONG HERED MUSC DYSTRPHY
 3591 HERED PROG MUSC DYSTRPHY
 3592 MYOTONIC DISORDERS
 35922 MYOTONIA CONGENITA
 35923 MYOTONIC CHONDRODYSTRPHY
 3593 FAMIL PERIODIC PARALYSIS
 3594 TOXIC MYOPATHY
 3595 MYOPATHY IN ENDOCRIN DIS
 3596 INFL MYOPATHY IN OTH DIS
 35981 CRITICAL ILLNESS MYOPATHY
 35989 MYOPATHIES NEC
 3599 MYOPATHY NOS

Exclude patients with selected ENT procedures

“Procedure Only”

27.31 Local excision or destruction of lesion or tissue of bony palate
 29.39 Other excision or destruction of lesion or tissue of pharynx
 29.4 Plastic operation on pharynx

 29.91 Dilation of pharynx
 30.09 Other excision or destruction of lesion or tissue of larynx
 30.22 Vocal cordectomy
 30.29 Other partial laryngectomy
 31.3 Other incision of larynx or trachea
 31.5 Local excision or destruction of lesion or tissue of trachea
 31.69 Other repair of larynx
 31.73 Closure of other fistula of trachea
 31.75 Reconstruction of trachea and construction of artificial larynx
 31.79 Other repair and plastic operations on trachea
 31.98 Other operations on larynx
 31.99 Other operations on trachea
 29.0 Pharyngotomy
 29.33 Pharyngectomy (partial)
 29.53 Closure of other fistula of pharynx
 29.59 Other repair of pharynx
 30.21 Epiglottidectomy
 30.3 Complete laryngectomy
 30.4 Radical laryngectomy
 25.3 Complete glossectomy
 25.4 Radical glossectomy

Exclude Patients with selected craniofacial procedures when accompanied by craniofacial abnormality diagnosis

25.2 Partial glossectomy
 25.59 Other repair and plastic operations on tongue
 27.62 Correction of cleft palate
 27.63 Revision of cleft palate repair
 27.69 Other plastic repair of palate
 29.31 Cricopharyngeal myotomy
 76.65 Segmental osteoplasty [osteotomy] of maxilla

	<p>76.66 Total osteoplasty [osteotomy] of maxilla 76.46 Other reconstruction of other facial bone 76.69 Other facial bone repair 76.91 Bone graft to facial bone 27.32 Wide excision or destruction of lesion or tissue of bony palate</p> <p>Craniofacial Dx</p> <p>744.83 Macrostomia 744.84 Microstomia 744.9 Unspecified anomalies of face and neck 748.3 Congenital anomalies of skull and face bones 756.0 Tracheomalacia and congenital tracheal stenosis</p>														
7 (2a, 2h)	<p>Stratification Do the measure specifications require the results to be stratified? No ▶ If "other" describe:</p> <p>Identification of stratification variable(s): None</p> <p>Stratification Details (Definitions, codes with description): None</p>														
8 (2a, 2e)	<p>Risk Adjustment Does the measure require risk adjustment to account for differences in patient severity before the onset of care? Yes ▶ If yes, Statistical Risk Model, see Variables ▶ Is there a separate proprietary owner of the risk model? No</p> <p>Identify Risk Adjustment Variables: Age, gender, modified CMS-DRG, AHRQ Comorbidity index in a logistic model with a hospital random effect</p> <p>Detailed risk model: attached <input checked="" type="checkbox"/> OR Web page URL: See also http://qualityindicators.ahrq.gov/downloads/psi/psi_covariates_v31.pdf</p>														
9 (2a)	<p>Type of Score: Rate/proportion Calculation Algorithm: attached <input checked="" type="checkbox"/> OR Web page URL: See also http://qualityindicators.ahrq.gov/downloads/psi/psi_guide_v31.pdf</p> <p>Interpretation of Score (<i>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</i>) Better quality = Lower score ▶ If "Other", please describe:</p>														
10 (2a, 4a, 4b)	<p>Identify the required data elements(e.g., primary diagnosis, lab values, vital signs): Age, gender, days from admission to procedure, MDC, DRG, principal and secondary diagnoses ICD-9-CM codes, principal and secondary procedure ICD-9-CM codes, admission type</p> <p>Data dictionary/code table attached <input type="checkbox"/> OR Web page URL: http://qualityindicators.ahrq.gov/downloads/psi/psi_sas_documentation_v32.pdf</p> <p>Data Quality (2a) <i>Check all that apply</i></p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> Data are captured from an authoritative/accurate source (e.g., lab values from laboratory personnel) <input checked="" type="checkbox"/> Data are coded using recognized data standards <input checked="" type="checkbox"/> Method of capturing data electronically fits the workflow of the authoritative source <input checked="" type="checkbox"/> Data are available in EHRs <input checked="" type="checkbox"/> Data are auditable 														
11 (2a, 4b)	<p>Data Source and Data Collection Methods <i>Identifies the data source(s) necessary to implement the measure specifications. Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Electronic Health/Medical Record</td> <td><input type="checkbox"/> Paper Medical Record</td> </tr> <tr> <td><input type="checkbox"/> Electronic Clinical Database, Name:</td> <td><input type="checkbox"/> Standardized clinical instrument, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Clinical Registry, Name:</td> <td><input type="checkbox"/> Standardized patient survey, Name:</td> </tr> <tr> <td><input checked="" type="checkbox"/> Electronic Claims</td> <td><input type="checkbox"/> Standardized clinician survey, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Pharmacy data</td> <td><input type="checkbox"/> Other, Describe:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Lab data</td> <td></td> </tr> <tr> <td><input type="checkbox"/> Electronic source - other, Describe:</td> <td>Instrument/survey attached <input type="checkbox"/> OR Web page URL:</td> </tr> </table>	<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record	<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:	<input type="checkbox"/> Electronic Clinical Registry, Name:	<input type="checkbox"/> Standardized patient survey, Name:	<input checked="" type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:	<input type="checkbox"/> Electronic Pharmacy data	<input type="checkbox"/> Other, Describe:	<input type="checkbox"/> Electronic Lab data		<input type="checkbox"/> Electronic source - other, Describe:	Instrument/survey attached <input type="checkbox"/> OR Web page URL:
<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record														
<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:														
<input type="checkbox"/> Electronic Clinical Registry, Name:	<input type="checkbox"/> Standardized patient survey, Name:														
<input checked="" type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:														
<input type="checkbox"/> Electronic Pharmacy data	<input type="checkbox"/> Other, Describe:														
<input type="checkbox"/> Electronic Lab data															
<input type="checkbox"/> Electronic source - other, Describe:	Instrument/survey attached <input type="checkbox"/> OR Web page URL:														

12 (2a)	<p>Sampling <i>If measure is based on a sample, provide instructions and guidance on sample size.</i> Minimum sample size: 0 Instructions: None</p>																		
13 (2a)	<p>Type of Measure: Outcome ▶ If "Other", please describe: ▶ If part of a composite or paired with another measure, please identify composite or paired measure It will be included in the Patient Safety for Selected Indicators composite if endorsed</p>																		
14 (2a)	<p>Unit of Measurement/Analysis <i>(Who or what is being measured) Check all that apply.</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Can be measured at all levels</td> <td><input type="checkbox"/> Integrated delivery system</td> </tr> <tr> <td><input type="checkbox"/> Individual clinician (e.g., physician, nurse)</td> <td><input type="checkbox"/> Health plan</td> </tr> <tr> <td><input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)</td> <td><input type="checkbox"/> Community/Population</td> </tr> <tr> <td><input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)</td> <td><input type="checkbox"/> Other <i>(Please describe):</i></td> </tr> </table>	<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system	<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input type="checkbox"/> Health plan	<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input type="checkbox"/> Community/Population	<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other <i>(Please describe):</i>										
<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system																		
<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input type="checkbox"/> Health plan																		
<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input type="checkbox"/> Community/Population																		
<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other <i>(Please describe):</i>																		
15 (2a)	<p>Applicable Care Settings <i>Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Can be used in all healthcare settings</td> <td><input type="checkbox"/> Hospice</td> </tr> <tr> <td><input type="checkbox"/> Ambulatory Care (office/clinic)</td> <td><input checked="" type="checkbox"/> Hospital</td> </tr> <tr> <td><input type="checkbox"/> Behavioral Healthcare</td> <td><input type="checkbox"/> Long term acute care hospital</td> </tr> <tr> <td><input type="checkbox"/> Community Healthcare</td> <td><input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)</td> </tr> <tr> <td><input type="checkbox"/> Dialysis Facility</td> <td><input type="checkbox"/> Prescription Drug Plan</td> </tr> <tr> <td><input type="checkbox"/> Emergency Department</td> <td><input type="checkbox"/> Rehabilitation Facility</td> </tr> <tr> <td><input type="checkbox"/> EMS emergency medical services</td> <td><input type="checkbox"/> Substance Use Treatment Program/Center</td> </tr> <tr> <td><input type="checkbox"/> Health Plan</td> <td><input type="checkbox"/> Other <i>(Please describe):</i></td> </tr> <tr> <td><input type="checkbox"/> Home Health</td> <td></td> </tr> </table>	<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice	<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital	<input type="checkbox"/> Behavioral Healthcare	<input type="checkbox"/> Long term acute care hospital	<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)	<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan	<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility	<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center	<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other <i>(Please describe):</i>	<input type="checkbox"/> Home Health	
<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice																		
<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital																		
<input type="checkbox"/> Behavioral Healthcare	<input type="checkbox"/> Long term acute care hospital																		
<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)																		
<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan																		
<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility																		
<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center																		
<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other <i>(Please describe):</i>																		
<input type="checkbox"/> Home Health																			
IMPORTANCE TO MEASURE AND REPORT																			
<p>Note: This is a threshold criterion. If a measure is not judged to be sufficiently important to measure and report, it will not be evaluated against the remaining criteria.</p>																			
16 (1a)	<p>Addresses a Specific National Priority Partners Goal <i>Enter the numbers of the specific goals related to this measure (see list of goals on last page):</i> 3.1, 6.1</p>																		
17 (1a)	<p>If not related to NPP goal, identify high impact aspect of healthcare (select one)</p> <p>Summary of Evidence: None</p> <p>Citations² for Evidence: None</p>																		
18 (1b)	<p>Opportunity for Improvement <i>Provide evidence that demonstrates considerable variation, or overall poor performance, across providers.</i></p> <p>Summary of Evidence: For the 2005 Nationwide Inpatient Sample from the AHRQ Healthcare Cost and Utilization Project, there was a rate of 10.869 per 1000 eligible discharges. Rates were very similar between teaching (10.965) and nonteaching (10.816) hospitals, but were substantially higher at small hospitals with 100-299 beds (12.129) than at medium-sized hospitals with 300-499 beds (10.965) or large hospitals with 500 or more beds (9.944). Rates were also substantially higher at private for-profit hospitals (13.206) than at private nonprofit hospitals (10.396), with public hospitals in between. Among the 18 hospitals participating in a University HealthSystem Consortium study of this indicator, estimated rates of "Postoperative Respiratory Failure" ranged from 2.3% to 29.2%. This variation suggests that there may be substantial opportunity for improved performance.</p> <p>Re Solucient (Foster D, March 17, 2008):If all hospitals had the same performance on postoperative respiratory failure as the 100 Top Hospitals based on National Benchmarks for Success, then 3,208 deaths and \$295.444 million in hospital costs could have been avoided in 2000-2005.</p>																		

² Citations can include, but are not limited to journal articles, reports, web pages (URLs).
 NQF Measure Submission Form, V3.0

	<p>Citations for Evidence: http://www.qualityindicators.ahrq.gov/downloads/psi/psi_provider_comparative_v31.pdf http://hcupnet.ahrq.gov/</p>						
<p>19 (1b)</p>	<p>Disparities <i>Provide evidence that demonstrates disparity in care/outcomes related to the measure focus among populations.</i></p> <p>Summary of Evidence: Higher rates have been demonstrated among Medicaid beneficiaries (relative to either Medicare enrollees or privately insured individuals) and among persons living in zip codes with the lowest quartile of median household income (11.840) than among persons living in zip codes with the highest quartile of median household income (9.874).</p> <p>HCUP Statistical Brief #53 by Russo CA et al. (June 2008). Risk-adjusted rate of postoperative respiratory failure was 1.21 in Blacks, 1.15 in Hispanics, and 1.14 in Asian-Pacific Islanders, relative to white patients in 2005.</p> <p>Citations for evidence: Http://hcupnet.ahrq.gov</p>						
<p>20 (1c)</p>	<p>If measuring an Outcome Describe relevance to the national health goal/priority, condition, population, and/or care being addressed: This indicator reflects a major complication with significant associated adverse health consequences for patients and efficiency/resources consequences for the system</p> <p>If not measuring an outcome, provide evidence supporting this measure topic and grade the strength of the evidence <i>Summarize the evidence (including citations to source) supporting the focus of the measure as follows:</i></p> <ul style="list-style-type: none"> • <u>Intermediate outcome</u> - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit. • <u>Process</u> - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s). • <u>Structure</u> - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit. • <u>Patient experience</u> - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public. • <u>Access</u> - evidence that an association exists between access to a health service and the outcomes of, or experience with, care. • <u>Efficiency</u>- demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality. <p>Type of Evidence <i>Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Evidence-based guideline</td> <td><input checked="" type="checkbox"/> Quantitative research studies</td> </tr> <tr> <td><input type="checkbox"/> Meta-analysis</td> <td><input checked="" type="checkbox"/> Qualitative research studies</td> </tr> <tr> <td><input type="checkbox"/> Systematic synthesis of research</td> <td><input type="checkbox"/> Other (<i>Please describe</i>):</td> </tr> </table> <p>Overall Grade for Strength of the Evidence³ (<i>Use the USPSTF system, or if different, also describe how it relates to the USPSTF system</i>):</p> <p>Summary of Evidence (<i>provide guideline information below</i>): Coding or Criterion Validity. Recent studies on the Patient Safety Indicator Algorithm found that most cases flagged by this indicator</p>	<input type="checkbox"/> Evidence-based guideline	<input checked="" type="checkbox"/> Quantitative research studies	<input type="checkbox"/> Meta-analysis	<input checked="" type="checkbox"/> Qualitative research studies	<input type="checkbox"/> Systematic synthesis of research	<input type="checkbox"/> Other (<i>Please describe</i>):
<input type="checkbox"/> Evidence-based guideline	<input checked="" type="checkbox"/> Quantitative research studies						
<input type="checkbox"/> Meta-analysis	<input checked="" type="checkbox"/> Qualitative research studies						
<input type="checkbox"/> Systematic synthesis of research	<input type="checkbox"/> Other (<i>Please describe</i>):						

³The strength of the body of evidence for the specific measure focus should be systematically assessed and rated, e.g., USPSTF grading system www.ahrq.gov/clinic/uspstmeth.htm: A - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. B - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. C - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. D - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. I - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

occur in-hospital (range 93%-100). One study found a somewhat lower rate of 74%.{Houchens, 2008 #326}{Naessens, 2007 #321}{Bahl, 2008 #328}

Several studies have assessed the criterion validity of this indicator. A study in the VA found a 67% sensitivity, and a 66% Positive Predictive Value for the current Respiratory Failure PSI algorithm, when compared with detailed clinical data abstracted for the National Surgical Quality Improvement Program. The likelihood ratio was 134, indicating that a flagged case was 134 times as likely to have the complication of interest than an unflagged case.{Romano, 2008 #329}

Recently, the University HealthSystem Consortium (UHC) conducted a Clinical Benchmarking Project based on this "Postoperative Respiratory Failure" indicator between September 2007 and May 2008. A total of 18 volunteer hospitals from 15 states submitted patient-level data for 692 complete cases that were flagged by the AHRQ software. The goals of this project were to: (1) evaluate the predictive value of the AHRQ Patient Safety Indicator in identifying cases with postoperative respiratory failure; (2) explore clinical aspects of the prevention, recognition, and screening for postoperative respiratory failure; (3) describe factors that identify patients at increased risk for postoperative respiratory failure; and (4) share successful strategies for safe management of patients at risk for postoperative respiratory failure. Participating hospitals retrospectively reviewed 40 discharges flagged by the AHRQ software that occurred on or prior to June 30, 2007, enrolling cases in reverse chronological order by discharge date to avoid selection bias. UHC trained abstractors at each hospital who conducted these reviews, using tools and guidelines developed by UHC staff. Overall, 93.2% (645/692) of cases were accurately flagged by the AHRQ software as having postoperative respiratory failure. This estimate represents a weighted average of predictive values of 85% among cases flagged using only the diagnosis codes in the AHRQ definition, and 95% among cases flagged using the procedure codes in the AHRQ definition, respectively. False positives were due to preoperative respiratory failure (2.7%), respiratory failure at admission (1.3%), exclusionary neuromuscular disorders (0.4%), and lack of clinical criteria for respiratory failure (2.3%). The most common reason in the last category was that mechanical ventilation was prolonged due to neurologic concerns or a need for continued airway protection. Nonetheless, the percentage of true positive cases was 84% or greater at every participating hospital. About 94.5% of true positive cases had procedures that were performed in the operating room, and 31.2% required a postoperative tracheostomy (indicating significant morbidity). In terms of outcomes, 76.7% (495/645) of patients confirmed as having "Postoperative Respiratory Failure" survived the target admission. Of these survivors, 7.7% (38/495) remained on ventilator support at discharge and 3.6% (18/495) had a related readmission within 30 days of discharge.

Predictive Validity. A study utilizing the 2000 Nationwide Inpatient Sample found that cases flagged by the Postoperative Respiratory Failure PSI had a 21.8% increase in mortality, an average of 9.1 days longer length of stay and average of \$53,500 more charges for that hospital stay. (Zhan and Miller, 2003) This study was replicated by Rivard et al., who found similar results from Veterans Health Administration hospitals (24.2% increase in mortality, 8.6 day increase in LOS, and \$39,745 excess charges). Solucient (now a division of MedStat Thomson) also replicated these findings using Medicare data, showing that patients who experience this complication have 17-fold higher odds of death (with 2,000 attributable deaths in the US), about 10 days of excess LOS (with about 80,000 attributable hospital days in the US), and nearly \$25,000 in excess hospital costs, relative to patients of the same age and gender, with the same DRG and comorbidities, who did not experience the complication. HealthGrades reported that this was the second most frequent non-mortality Patient Safety Indicator among Medicare patients in 2004-2006, with an attributable cost of \$1.84 billion (HealthGrades 5th Annual Patient Safety in American Hospitals Study, 2008).

Citations for Evidence:

1. Geraci JM, Ashton CM, Kuykendall DH, Johnson ML, Wu L. International Classification of Diseases, 9th Revision, Clinical Modification codes in discharge abstracts are poor measures of complication occurrence in medical inpatients. *Med Care.* 1997;35(6):589-602.
2. Romano P. Can Administrative Data be Used to Ascertain Clinically Significant Postoperative Complications. *American Journal of Medical Quality.* Press.
3. Best W, Khuri S, Phelan M, et al. Identifying Patient Preoperative Risk Factors and Postoperative Adverse Events in Administrative Databases: Results from the Department of Veterans Affairs National Surgical Quality Improvement Program. *J Am Coll Surg.* 2002;194(3):257-266.
4. Hannan EL, Bernard HR, O'Donnell JF, Kilburn H, Jr. A methodology for targeting hospital cases for

	<p>quality of care record reviews. Am J Public Health. 1989;79(4):430-6.</p> <p>5. Needleman J, Buerhaus PI, Mattke S, Stewart M, Zelevinsky K. (Health Resources Services Administration). Nurse Staffing and Patient Outcomes in Hospitals. 2001 February 28. Report No.: 230-99-0021.</p>
21 (1c)	<p>Clinical Practice Guideline <i>Cite the guideline reference; quote the specific guideline recommendation related to the measure and the guideline author's assessment of the strength of the evidence; and summarize the rationale for using this guideline over others.</i></p> <p>Guideline Citation: None</p> <p>Specific guideline recommendation: None</p> <p>Guideline author's rating of strength of evidence (If different from USPSTF, also describe it and how it relates to USPSTF): None</p> <p>Rationale for using this guideline over others: None</p>
22 (1c)	<p>Controversy/Contradictory Evidence <i>Summarize any areas of controversy, contradictory evidence, or contradictory guidelines and provide citations.</i></p> <p>Summary: None</p> <p>Citations: None</p>
23 (1)	<p>Briefly describe how this measure (as specified) will facilitate significant gains in healthcare quality related to the specific priority goals and quality problems identified above: This complication results in significant increases in resource use and increases potential harm to patient. Recent validity studies have found that this indicator has a positive predictive validity. While less evidence is available regarding the preventability of this complication, it is clear that this complication results in less efficiency in care. Some potential causes of respiratory failure are associated with processes of care.</p>
SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
<p>Note: Testing and results should be summarized in this form. However, additional detail and reports may be submitted as supplemental information or provided as a web page URL. If a measure has not been tested, it is only potentially eligible for time-limited endorsement.</p>	
24	<p>Supplemental Testing Information: attached <input type="checkbox"/> OR Web page URL: http://www.ahrq.gov/downloads/pub/evidence/pdf/psi/psi.pdf</p>
25 (2b)	<p>Reliability Testing</p> <p>Data/sample: See summary of studies conducted above</p> <p>Analytic Method: See above</p> <p>Testing Results: See above</p>
26 (2c)	<p>Validity Testing</p> <p>Data/sample: See summary of studies conducted above</p> <p>Analytic Method: See above</p> <p>Testing Results: See above</p>
27 (2d)	<p>Measure Exclusions <i>Provide evidence to justify exclusion(s) and analysis of impact on measure results during testing.</i></p> <p>Summary of Evidence supporting exclusion(s):</p> <p>a. The AHRQ QI support team underwent analyses to test excluding patients with specific craniofacial, laryngeal and pharyngeal surgery.</p>

	<p>b. A VA study examined the evidence for including procedure codes for delayed extubation or reintubation to improve sensitivity of the indicator.</p> <p>Citations for Evidence: Internal studies</p> <p>Data/sample: a. 2003 Nationwide Inpatient Sample b. Romano PS, Mull HJ, Rivard PE, Zhao S, Henderson WG, Loveland S, Tsilimingras D, Christiansen CL, Rosen AK. Validity of Selected AHRQ Patient Safety Indicators Based on VA National Surgical Quality Improvement Program Data. Health Serv Res. 2008 Sep 17</p> <p>Analytic Method: a. We calculated risk ratios for patients with any procedure code affecting the larynx, pharynx, and craniofacial area that may result in prolonged intubation and thus trigger the indicator without true respiratory failure b. Compared sensitivity and specificity of the indicator using NSQUIP as the gold standard before and after adding the definitional change.</p> <p>Testing Results: a. The procedures listed in the procedure exclusion list had higher risk ratios b. adding procedure codes increased the sensitivity of the indicator 17% to 63% without significant decreases in specificity.</p>
<p>28 (2e)</p>	<p>Risk Adjustment Testing Summarize the testing used to determine the need (or no need) for risk adjustment and the statistical performance of the risk adjustment method.</p> <p>Data/sample: Hospital discharge data from 2002-2004 for states participating in the HCUP State Inpatient Database (33-38 states representing approximately 30 adult million discharges and 4,500 hospitals per year)</p> <p>Analytic Method: AHRQ convened a Risk Adjustment and Hierarchical Modeling workgroup to review the AHRQ QI risk adjustment methodology and to recommend analyses to support adoption of the hierarchical model (a logistic with a hospital random effect) which the workgroup recommended and is currently in use.</p> <p>Testing Results: See attached report from the AHRQ Risk Adjustment and Hierarchical Modeling Workgroup</p> <p>► If outcome or resource use measure not risk adjusted, provide rationale: None</p>
<p>29 (2g)</p>	<p>Testing comparability of results when more than 1 data method is specified (e.g., administrative claims or chart abstraction)</p> <p>Data/sample: None</p> <p>Analytic Method: None</p> <p>Results: None</p>
<p>30 (2f)</p>	<p>Provide Measure Results from Testing or Current Use Results from testing</p> <p>Data/sample: Hospital discharge data from 2002-2004 for states participating in the HCUP State Inpatient Database (33-38 states representing approximately 30 adult million discharges and 4,500 hospitals per year)</p> <p>Methods to identify statistically significant and practically/meaningfully differences in performance: We use an empirical bayes univariate shrinkage estimator to compute the "signal ratio" (a measure of reliability) and estimate the amount of true underlying "signal" (a measure of variation) in the observed distribution of hospital performance</p> <p>Results: We find a signal ratio of 0.8664 (which is considered acceptable) and a signal standard deviation of 3.687 on an average hospital rate of 8.404, meaning that 95% of hospitals fall within the range of 1.03 and 15.77 per 1,000 discharges.</p>
<p>31</p>	<p>Identification of Disparities</p> <p>► If measure is stratified by factors related to disparities (i.e. race/ethnicity, primary language, gender,</p>

(2h)	SES, health literacy), provide stratified results: None ► If disparities have been reported/identified, but measure is not specified to detect disparities, provide rationale: None
USABILITY	
32	Current Use In use If in use, how widely used Nationally ► If "other," please describe:
(3)	<input checked="" type="checkbox"/> Used in a public reporting initiative, name of initiative: Multiple states Sample report attached <input checked="" type="checkbox"/> OR Web page URL:
33	Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)
(3a)	<p>Data/sample: A research team from Weill Cornell Medical College's Department of Public Health and the School of Public Affairs, Baruch College, has developed the attached Hospital Quality Model Report for the Agency for Healthcare Research & Quality (AHRQ).</p> <p>Methods: This Model Report is based on: 1) Extensive search and analysis of the literature on hospital quality measurement and reporting, as well as public reporting on health care quality more broadly; 2) Interviews with experts, purchasers, staff of purchasing coalitions, and executives of integrated health care delivery systems who were responsible for quality in their facilities; 2) Two focus groups with chief medical officers of hospitals and/or systems and two focus groups with quality managers from a broad mix of hospitals; 3) Four focus groups with members of the public who had recently experienced a hospital admission; and 4) Two rounds of cognitive interviews (a total of 19 interviews) to test draft versions of the Model Report with members of the public with recent hospital experience, who had basic computer literacy but widely varying levels of education.</p> <p>Results: The model report was assessed in the NQF report National Voluntary Consensus Standards for Hospital Care: Additional Priorities--2007, Part 3: Guidelines for Consumer-focused Public Reporting available at http://www.qualityforum.org/projects/ongoing/hosp-priorities2007.</p>
34	Relation to other NQF-endorsed™ measures
(3b, 3c)	► Is this measure similar or related to measure(s) already endorsed by NQF (on the same topic or the same target population)? <i>Measures can be found at www.qualityforum.org under Core Documents.</i> Check all that apply <input type="checkbox"/> Have not looked at other NQF measures <input type="checkbox"/> Other measure(s) on same topic <input type="checkbox"/> Other measure(s) for same target population <input checked="" type="checkbox"/> No similar or related measures Name of similar or related NQF-endorsed™ measure(s): None Are the measure specifications harmonized with existing NQF-endorsed™ measures? (select one) ► If not fully harmonized, provide rationale: None Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: None
FEASIBILITY	
35	How are the required data elements generated? <i>Check all that apply</i>
(4a)	<input type="checkbox"/> Data elements are generated concurrent with and as a byproduct of care processes during care delivery (e.g., blood pressure or other assessment recorded by personnel conducting the assessment) <input type="checkbox"/> Data elements are generated from a patient survey (e.g., CAHPS) <input checked="" type="checkbox"/> Data elements are generated through coding performed by someone other than the person who obtained the original information (e.g., DRG or ICD-9 coding on claims) <input type="checkbox"/> Other, Please describe:
36	Electronic Sources All data elements

<p>(4b)</p>	<p>► If all data elements are not in electronic sources, specify the near-term path to electronic collection by most providers:</p> <p>► Specify the data elements for the electronic health record: Required data elements: Age in years at admission; International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) principal and secondary diagnosis codes; principal and secondary procedure codes; Diagnosis Related Group (DRGs) and Major Diagnostic Category (MDC); dates of major operating room procedures (in relation to admission date), admission type.</p> <p>For details on data element definitions and allowable values, see PSI SAS Documentation, p. 11 (Table 4. Data Elements and Coding Conventions). http://www.qualityindicators.ahrq.gov/downloads/psi/psi_sas_documentation_v31.pdf</p>
<p>37 (4c)</p>	<p>Do the specified exclusions require additional data sources beyond what is required for the other specifications? No</p> <p>► If yes, provide justification: None</p>
<p>38 (4d)</p>	<p>Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure: This indicator has been shown to have a low false positive rate in studies, but this indicator is dependent on physician documentation.</p> <p>Describe how could these potential problems be audited: Periodic chart review, although given the definition reliant on procedure codes, such review is likely unnecessary</p> <p>Did you audit for these potential problems during testing? No If yes, provide results:</p>
<p>39 (4e)</p>	<p>Testing feasibility Describe what have you learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: None</p>
<p>CONTACT INFORMATION</p>	
<p>40</p>	<p>Web Page URL for Measure Information Describe where users (implementers) should go for more details on specifications of measures, or assistance in implementing the measure. Web page URL: http://qualityindicators.ahrq.gov</p>
<p>41</p>	<p>Measure Intellectual Property Agreement Owner Point of Contact First Name: Mamatha MI: Last Name: Pancholi Credentials (MD, MPH, etc.): Organization: Agency for Healthcare Research and Quality Street Address: 540 Gaither Road City: Rockville State: MD ZIP: 20850 Email: mamatha.pancholi@ahrq.hhs.gov Telephone: 301-427-1470 ext:</p>
<p>42</p>	<p>Measure Submission Point of Contact If different than IP Owner Contact First Name: MI: Last Name: Credentials (MD, MPH, etc.): Organization: Street Address: City: State: ZIP: Email: Telephone: ext:</p>
<p>43</p>	<p>Measure Developer Point of Contact If different than IP Owner Contact First Name: MI: Last Name: Credentials (MD, MPH, etc.): Organization: Street Address: City: State: ZIP: Email: Telephone: ext:</p>
<p>44</p>	<p>Measure Steward Point of Contact If different than IP Owner Contact <i>Identifies the organization that will take responsibility for updating the measure and assuring it is consistent with the scientific evidence and current coding schema; the steward of the measure may be different than the developer.</i> First Name: MI: Last Name: Credentials (MD, MPH, etc.):</p>

	<p>Organization: Street Address: City: State: ZIP: Email: Telephone: ext</p>
ADDITIONAL INFORMATION	
<p>45</p>	<p>Workgroup/Expert Panel involved in measure development <i>Workgroup/panel used</i></p> <p>▶ If workgroup used, describe the members' role in measure development: <i>Clinical Panel provided input on the definition of the indicator and rated the usefulness of the indicator for quality improvement purposes. Additional information on the panel review and results can be found at: http://www.qualityindicators.ahrq.gov/downloads/technical/psi_technical_review.zip.</i></p> <p>▶ Provide a list of workgroup/panel members' names and organizations:</p> <p><i>Robert Kozol, MD, MSA, Surgeon Farmington, CT University of Connecticut Nominated by the American College of Surgeons</i></p> <p><i>Steven Liu, MD, Hospitalist Atlanta, GA Emory University School of Medicine Nominated by the National Association of Inpatient Physicians</i></p> <p><i>Lenora Maze, MSN, Critical care nurse Indianapolis, IN Wishard Health Services Nominated by the Substitute for American Association of Critical-Care Nurses Nominee</i></p> <p><i>Valerie Palda, MD, MSc, Internist Toronto, ON University of Toronto Nominated by the American College of Physicians</i></p> <p><i>Sanjay Saint, MD, MPH, Hospitalist Ann Arbor, MI University of Michigan Medical School Nominated by the National Association of Inpatient Physicians</i></p> <p><i>Patrice Spera, RN, MS, Perioperative nurse Seminole, FL Tampa General Hospital Nominated by the Association of Peri-Operative Registered Nurses</i></p> <p><i>Joseph Basler, MD, PhD, Urologist San Antonio, TX University of Texas Health Science Center Nominated by the American Urologic Association</i></p> <p><i>John Fung, MD, Transplant surgeon Pittsburgh, PA University of Pittsburgh Nominated by the American Society of Transplant Surgeons</i></p> <p><i>Charles Kenny, MD, Orthopedic surgeon Stockbridge, MA Fairview Hospital Nominated by the American Academy of Orthopedic Surgeons</i></p> <p><i>John Kestle, MD, MSc, Pediatric neurosurgeon</i></p>

	<p>Salt Lake City, UT University of Utah Nominated by the American Association of Neurological Surgeons</p> <p>Michael Klassen, MD, Joint and arthroscopic surgeon Monterey, CA Community Hospital of the Monterey Peninsula Nominated by the American Academy of Orthopedic Surgeons</p> <p>George Lucas, MD, Orthopedic surgeon - hand surgery Wichita, KS University of Kansas, Wichita Nominated by the American Academy of Hand Surgeon</p> <p>Dennis Maiman, MD, PhD, Neurosurgeon- spine surgery Milwaukee, WI Froedert Memorial Lutheran Hospital Nominated by the North American Spine Society</p> <p>Richard Nelson, MD, Colon and rectal surgeon Chicago, IL University of Illinois Nominated by the American Society of Colon and Rectal Surgeons</p> <p>Michael Stamos, MD, Colon and rectal surgeon Torrance, CA University of California - Los Angeles School of Medicine Nominated by the American College of Surgeons</p>
46	<p>Measure Developer/Steward Updates and Ongoing Maintenance Year the measure was first released: 2003 Month and Year of most recent revision: March 2008 What is the frequency for review/update of this measure? Annually When is the next scheduled review/update for this measure? Spring 2009</p>
47	<p>Copyright statement/disclaimers: None</p>
48	<p>Additional Information: None</p>
49	<p>I have checked that the submission is complete and any blank fields indicate that no information is provided. <input checked="" type="checkbox"/></p>
50	<p>Date of Submission (MM/DD/YY): 11/21/2008</p>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.0

August 2008

The measure information you submit will be shared with NQF’s Steering Committees and Technical Advisory Panels to evaluate measures against the NQF criteria of importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Four conditions (as indicated below) must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, importance to measure and report, in order to be evaluated against the remaining criteria. References to the specific measure evaluation criteria are provided in parentheses following the item numbers. Please refer to the *Measure Evaluation Criteria* for more information at www.qualityforum.org under Core Documents. Additional guidance is being developed and when available will be posted on the NQF website.

Use the tab or arrow (↓→) keys to move the cursor to the next field (or back ←↑). There are three types of response fields:

- drop-down menus - select one response;
- check boxes - check as many as apply; and
- text fields - you can copy and paste text into these fields or enter text; these fields are not limited in size, but in most cases, we ask that you summarize the requested information.

Please note that URL hyperlinks do not work in the form; you will need to type them into your web browser.

Be sure to answer all questions. Fields that are left blank will be interpreted as no or none. Information must be provided in this form. Attachments are not allowed except when specifically requested or to provide additional detail or source documents for information that is summarized in this form. If you have important information that is not addressed by the questions, they can be entered into item #48 near the end of the form.

For questions about this form, please contact the NQF Project Director listed in the corresponding call for measures.

CONDITIONS FOR CONSIDERATION BY NQF	
	Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards.
A (A)	Public domain or Intellectual Property Agreement signed: Public domain - IP agreement not required (If no, do not submit) <i>Template for the Intellectual Property Agreement is available at www.qualityforum.org under Core Documents.</i>
B (B)	Measure steward/maintenance: Is there an identified responsible entity and process to maintain and update the measure on a schedule commensurate with clinical innovation, but at least every 3 years? Yes, information provided in contact section (If no, do not submit)
C (C)	Intended use: Does the intended use of the measure include BOTH public reporting AND quality improvement? Yes (If no, do not submit)
D (D)	Fully developed and tested: Is the measure fully developed AND tested? No, testing will be completed within 24 months (If not tested and no plans for testing within 24 months, do not submit)

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.0

August 2008

	(for NQF staff use) NQF Review #: HOE-008-08	NQF Project: Hospital Outcomes and Efficiency
MEASURE SPECIFICATIONS & DESCRIPTIVE INFORMATION		
1	Information current as of (date- MM/DD/YY): 11-14-08	
2	Title of Measure: Hospital specific risk-adjusted measure of mortality or one or more major complications within 30 days of a lower extremity bypass (LEB).	
3	Brief description of measure ¹ : Hospital specific risk-adjusted measure of mortality or one or more of the following major complications (cardiac arrest, myocardial infarction, CVA/stroke, on ventilator >48 hours, acute renal failure (requiring dialysis), bleeding/transfusions, graft/prosthesis/flap failure, septic shock, sepsis, and organ space surgical site infection), within 30 days of a lower extremity bypass (LEB).	
4 (2a)	<p>Numerator Statement: Note: This outcome measure does not have a traditional numerator and denominator like CMS core process measure; thus, we use this field to define our statistically-adjusted outcome measure.</p> <p>Hierarchical logistic regression modeling was used to calculate a hospital-specific lower extremity bypass standardized outcome ratio (LEBSOR). This is calculated as the ratio of “predicted” number of outcomes to the “expected” number of outcomes. For each hospital, the “numerator” of the ratio component of the LEBSOR is the predicted number of deaths or major complications within 30 days of LEB surgery given the hospital’s performance with its observed case mix. The “denominator” is the expected number of death and major complications given the average of all hospital’s case mix effects. By convention, the term “predicted” describes the numerator result, which is calculated using the hospital-specific intercept term. The “expected” is used for the denominator, which is calculated using the average hospital intercept term. Operationally, the expected number of death and major complications for each hospital is obtained by regressing the risk factors (see #16) on the complications using all hospitals in our sample, applying the subsequent estimated regression coefficients to the patient characteristics observed in the hospital, adding the average of the hospital-specific intercepts, transforming, and then summing over all patients in the hospital to get a value. This is a form of indirect standardization. The predicted hospital outcome is the number of deaths and major complications estimated in the “specific” hospital given its performance and case mix. Operationally, this is accomplished by estimating a hospital-specific intercept that herein represents baseline complications risk within the hospital, applying the estimated regression coefficients to the patient characteristics in the hospital, transforming, and then summing over all patients in the hospital to get a value.</p> <p>Time Window: For development, 3 years of data (July 2004- June 2007). For public reporting, the timeframe has not been determined.</p> <p>Numerator Details (Definitions, codes with description):</p>	
5 (2a)	<p>Denominator Statement: The measure is risk adjusted, please see Section # 8 below.</p> <p>Time Window: For development, 3 years of data (July 2004- June 2007). For public reporting, the timeframe has not been determined.</p> <p>Denominator Details (Definitions, codes with description): We are using this field to specify the codes that define the LEB patient cohort. 35537 - Bypass graft, with vein; aortoiliac 35538 - Bypass graft, with vein; aortobi-iliac 35539 - Bypass graft, with vein; aortofemoral</p>	

¹ Example of measure description: Percentage of adult patients with diabetes aged 18-75 years receiving one or more A1c test(s) per year.
 NQF Measure Submission Form, V3.0

	<p>35540 - Bypass graft, with vein; aortobifemoral 35541 - Bypass graft with vein, aortoiliac or bi-iliac 35546 - Bypass graft with vein, aortofemoral or bifemoral 35548 - Bypass graft, with vein; aortoiliofemoral, unilateral 35549 - Bypass graft, with vein; aortoiliofemoral, bilateral 35551 - Bypass graft, with vein; aortofemoral-popliteal 35556 - Bypass graft, with vein; femoral-popliteal 35558 - Bypass graft, with vein; femoral-femoral, 35563 - Bypass graft, with vein; ilioiliac, 35565 - Bypass graft, with vein; iliofemoral, 35566 - Bypass graft, with vein; femoral-anterior tibial, posterior tibial, peroneal artery or other distal vessels 35571 - Bypass graft, with vein; popliteal-tibial, -peroneal artery or other distal vessels 35583 - In-situ vein bypass; femoral-popliteal 35585 - In-situ vein bypass; femoral-anterior tibial, posterior tibial, or peroneal artery 35587 - Bypass graft, with vein; femoral-femoral 35623 - Bypass graft, with other than vein; axillary-popliteal or -tibial 35637 - Bypass graft, with other than vein; aortoiliac 35638 - Bypass graft, with other than vein; aortobi-iliac 35646 - Bypass graft, with other than vein; aortobifemoral 35647 - Bypass graft, with other than vein; aortofemoral 35651 - Bypass graft, with other than vein; aortofemoral-popliteal 35654 - Bypass graft, with other than vein; axillary-femoral-femoral 35656 - Bypass graft, with other than vein; femoral-popliteal 35661 - Bypass graft, with other than vein; femoral-femoral 35663 - Bypass graft, with other than vein; ilioiliac 35665 - Bypass graft, with other than vein; iliofemoral 35666 - Bypass graft, with other than vein; femoral-anterior tibial, posterior tibial, or peroneal artery 35671 - Bypass graft, with other than vein; popliteal-tibial or -peroneal artery 35700 - Reoperation, femoral-popliteal or femoral (popliteal)-anterior tibial, posterior tibial, peroneal artery, or other distal vessels, more than one month after original operation (List separately in addition to code for primary procedure) 35721 - Exploration (not followed by surgical repair), with or without lysis of artery; femoral artery 35741 - Exploration (not followed by surgical repair), with or without lysis of artery; popliteal artery 35879 - Revision, lower extremity arterial bypass, without thrombectomy, open; with vein patch angioplasty 35881 - Revision, lower extremity arterial bypass, without thrombectomy, open; with segmental vein interposition 35883 - Revision, femoral anastomosis of synthetic arterial bypass graft in groin, open; with nonautogenous patch graft (eg, Dacron, ePTFE, bovine pericardium) 35884 - Revision, femoral anastomosis of synthetic arterial bypass graft in groin, open; with autogenous vein patch graft</p>
<p>6 (2a, 2d)</p>	<p>Denominator Exclusions: No exclusions Denominator Exclusion Details (Definitions, codes with description): N/A</p>
<p>7 (2a, 2h)</p>	<p>Stratification Do the measure specifications require the results to be stratified? No ▶ If "other" describe: N/A Identification of stratification variable(s): N/A Stratification Details (Definitions, codes with description): N/A</p>
<p>8 (2a, 2e)</p>	<p>Risk Adjustment Does the measure require risk adjustment to account for differences in patient severity before the onset of care? Yes ▶ If yes, Statistical Risk Model, see Variables ▶ Is there a separate proprietary owner of the risk model? No</p>

	<p>Identify Risk Adjustment Variables:</p> <ol style="list-style-type: none"> 1. FUNCTIONAL STATUS: This variable focuses on the patient’s abilities to perform activities of daily living (ADLs) in the 30 days prior to surgery. Activities of daily living are defined as ‘the activities usually performed in the course of a normal day in a person’s life’. ADLs include: bathing, feeding, dressing, toileting, and mobility. Report the corresponding level of self-care for activities of daily living demonstrated by this patient for the following two time points: (a) prior to the current illness, and (b) at the time the patient is being considered as a candidate for surgery (which should be no longer than 30 days prior to surgery). If the patient’s status changes prior to surgery, that change should be reflected in your assessment of (b). For each of these time points, report the level of functional health status as defined by the following criteria. 1) Independent: The patient does not require assistance from another person for any activities of daily living. This includes a person who is able to function independently with prosthetics, equipment, or devices; 2) Partially dependent: The patient requires some assistance from another person for activities of daily living. This includes a person who utilizes prosthetics, equipment, or devices but still requires some assistance from another person for ADLs; 3) Totally dependent: The patient requires total assistance for all activities of daily living. 2. EMERGENCY SURGERY: An emergency case is usually performed as soon as possible and no later than 12 hours after the patient has been admitted to the hospital or after the onset of related preoperative symptomatology. Answer ‘yes’ if the surgeon and anesthesiologist report the case as emergent. 3. WORK RVU: Relative Value Unit: a factor tied to CPT codes developed and maintained by CMS, which is used in pricing of medical services 4. SGOT > 40: Pre-operative Lab Value 5. SERUM ALBUMIN: Pre-operative Lab Value 6. ASA CLASS: American Society of Anesthesiology class: Class I. Normal healthy patient; Class II. Patient with mild systemic disease Class III. Patient with severe systemic disease; Class IV. Patient with severe systemic disease that is a constant threat to life; Class V. a moribund patient who is not expected to survive without the operation 7. REST PAIN/GANGRENE: Rest pain is a more severe form of ischemic pain due to occlusive disease, which occurs at rest and is manifested as a severe, unrelenting pain aggravated by elevation and often preventing sleep. Gangrene is a marked skin discoloration and disruption indicative of death and decay of tissues in the extremities due to severe and prolonged ischemia. Include patients with ischemic ulceration and/or tissue loss related to peripheral vascular disease. Do not include Fournier’s gangrene. 8. TRANSFUSION >4 units within 72 hours of surgery: Preoperative loss of blood necessitating a minimum of 5 units of whole blood/packed red cells transfused during the 72 hours prior to surgery including any blood transfused in the emergency room. 9. MALE: Gender 10. CREATININE > 1.2 mg/dl: Pre-operative Lab Value <p>Detailed risk model: attached <input checked="" type="checkbox"/> OR Web page URL:</p>
<p>9 (2a)</p>	<p>Type of Score: Ratio Calculation Algorithm: attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p>Interpretation of Score (<i>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</i>) Other ► If “Other”, please describe: For each hospital, a 30-day all-cause lower extremity bypass standardized outcome ratio (LEBSOR) and an interval estimate for the LEBSOR were calculated, which expresses the level of uncertainty around the point estimate. The Risk Standardized Complication Rate</p>

	(RSCR) and interval estimate can be used to classify hospital performance (e.g., higher than expected, as expected, or lower than expected) and to compare hospitals (e.g., nationally, regionally, within peer groups).																		
10 (2a, 4a, 4b)	<p>Identify the required data elements(e.g., primary diagnosis, lab values, vital signs): see appendix in final report</p> <p>Data dictionary/code table attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p>Data Quality (2a) Check all that apply</p> <p><input checked="" type="checkbox"/> Data are captured from an authoritative/accurate source (e.g., lab values from laboratory personnel)</p> <p><input checked="" type="checkbox"/> Data are coded using recognized data standards</p> <p><input checked="" type="checkbox"/> Method of capturing data electronically fits the workflow of the authoritative source</p> <p><input checked="" type="checkbox"/> Data are available in EHRs</p> <p><input checked="" type="checkbox"/> Data are auditable</p>																		
11 (2a, 4b)	<p>Data Source and Data Collection Methods <i>Identifies the data source(s) necessary to implement the measure specifications. Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Electronic Health/Medical Record</td> <td><input type="checkbox"/> Paper Medical Record</td> </tr> <tr> <td><input type="checkbox"/> Electronic Clinical Database, Name:</td> <td><input type="checkbox"/> Standardized clinical instrument, Name:</td> </tr> <tr> <td><input checked="" type="checkbox"/> Electronic Clinical Registry, Name: NSQIP, or any other alternative data collection method</td> <td><input type="checkbox"/> Standardized patient survey, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Claims</td> <td><input type="checkbox"/> Standardized clinician survey, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Pharmacy data</td> <td><input type="checkbox"/> Other, Describe:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Lab data</td> <td></td> </tr> <tr> <td><input type="checkbox"/> Electronic source - other, Describe:</td> <td></td> </tr> </table> <p>Instrument/survey attached <input type="checkbox"/> OR Web page URL:</p>	<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record	<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:	<input checked="" type="checkbox"/> Electronic Clinical Registry, Name: NSQIP, or any other alternative data collection method	<input type="checkbox"/> Standardized patient survey, Name:	<input type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:	<input type="checkbox"/> Electronic Pharmacy data	<input type="checkbox"/> Other, Describe:	<input type="checkbox"/> Electronic Lab data		<input type="checkbox"/> Electronic source - other, Describe:					
<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record																		
<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:																		
<input checked="" type="checkbox"/> Electronic Clinical Registry, Name: NSQIP, or any other alternative data collection method	<input type="checkbox"/> Standardized patient survey, Name:																		
<input type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:																		
<input type="checkbox"/> Electronic Pharmacy data	<input type="checkbox"/> Other, Describe:																		
<input type="checkbox"/> Electronic Lab data																			
<input type="checkbox"/> Electronic source - other, Describe:																			
12 (2a)	<p>Sampling <i>If measure is based on a sample, provide instructions and guidance on sample size.</i></p> <p>Minimum sample size: 0</p> <p>Instructions: For public reporting, sampling strategies and minimum sample size have not been determined</p>																		
13 (2a)	<p>Type of Measure: Outcome ▶ If "Other", please describe:</p> <p>▶ If part of a composite or paired with another measure, please identify composite or paired measure N/A</p>																		
14 (2a)	<p>Unit of Measurement/Analysis <i>(Who or what is being measured) Check all that apply.</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Can be measured at all levels</td> <td><input type="checkbox"/> Integrated delivery system</td> </tr> <tr> <td><input type="checkbox"/> Individual clinician (e.g., physician, nurse)</td> <td><input type="checkbox"/> Health plan</td> </tr> <tr> <td><input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)</td> <td><input type="checkbox"/> Community/Population</td> </tr> <tr> <td><input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)</td> <td><input type="checkbox"/> Other <i>(Please describe):</i></td> </tr> </table>	<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system	<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input type="checkbox"/> Health plan	<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input type="checkbox"/> Community/Population	<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other <i>(Please describe):</i>										
<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system																		
<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input type="checkbox"/> Health plan																		
<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input type="checkbox"/> Community/Population																		
<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other <i>(Please describe):</i>																		
15 (2a)	<p>Applicable Care Settings <i>Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Can be used in all healthcare settings</td> <td><input type="checkbox"/> Hospice</td> </tr> <tr> <td><input type="checkbox"/> Ambulatory Care (office/clinic)</td> <td><input checked="" type="checkbox"/> Hospital</td> </tr> <tr> <td><input type="checkbox"/> Behavioral Healthcare</td> <td><input checked="" type="checkbox"/> Long term acute care hospital</td> </tr> <tr> <td><input type="checkbox"/> Community Healthcare</td> <td><input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)</td> </tr> <tr> <td><input type="checkbox"/> Dialysis Facility</td> <td><input type="checkbox"/> Prescription Drug Plan</td> </tr> <tr> <td><input type="checkbox"/> Emergency Department</td> <td><input type="checkbox"/> Rehabilitation Facility</td> </tr> <tr> <td><input type="checkbox"/> EMS emergency medical services</td> <td><input type="checkbox"/> Substance Use Treatment Program/Center</td> </tr> <tr> <td><input type="checkbox"/> Health Plan</td> <td><input type="checkbox"/> Other <i>(Please describe):</i></td> </tr> <tr> <td><input type="checkbox"/> Home Health</td> <td></td> </tr> </table>	<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice	<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital	<input type="checkbox"/> Behavioral Healthcare	<input checked="" type="checkbox"/> Long term acute care hospital	<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)	<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan	<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility	<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center	<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other <i>(Please describe):</i>	<input type="checkbox"/> Home Health	
<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice																		
<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital																		
<input type="checkbox"/> Behavioral Healthcare	<input checked="" type="checkbox"/> Long term acute care hospital																		
<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)																		
<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan																		
<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility																		
<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center																		
<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other <i>(Please describe):</i>																		
<input type="checkbox"/> Home Health																			
IMPORTANCE TO MEASURE AND REPORT																			
Note: This is a threshold criterion. If a measure is not judged to be sufficiently important to measure and report, it will not be evaluated against the remaining criteria.																			
16 (1a)	<p>Addresses a Specific National Priority Partners Goal <i>Enter the numbers of the specific goals related to this measure (see list of goals on last page):</i> 3.1, 3.3, 3.4</p>																		

<p>17 (1a)</p>	<p>If not related to NPP goal, identify high impact aspect of healthcare (select one)</p> <p>Summary of Evidence: N/A</p> <p>Citations² for Evidence: N/A</p>
<p>18 (1b)</p>	<p>Opportunity for Improvement <i>Provide evidence that demonstrates considerable variation, or overall poor performance, across providers.</i></p> <p>Summary of Evidence: The composite outcome occurs at 16.9% for LEB patients which indicates that a considerable proportion of LEB patients either die or experience a major complication. This indicates a gap in quality and represents room for improvement.</p> <p>Citations for Evidence: N/A</p>
<p>19 (1b)</p>	<p>Disparities <i>Provide evidence that demonstrates disparity in care/outcomes related to the measure focus among populations.</i></p> <p>Summary of Evidence: We have not examined health disparities associated with this measure.</p> <p>Citations for evidence: N/A</p>
<p>20 (1c)</p>	<p>If measuring an Outcome Describe relevance to the national health goal/priority, condition, population, and/or care being addressed:</p> <p>Lower extremity vascular disease is common and will increase with the increasing age of the U. S. population. Bypass surgery is the predominant treatment modality for this disease, and is therefore an important procedure for which quality should be assessed and continually improved upon. Measuring and tracking mortality and major complications experienced by LEB patients will help minimize potential adverse outcomes of LEB, which can be costly to our healthcare system and are important to the patient. Using such measure will help drive system-wide improvement that will result in reducing the incidence of healthcare-associated infections to zero and reducing premature and preventable mortality to best in class. Additionally, tracking and acting on this measure will encourage more accurate and complete medication reconciliation across the continuum of care, reduce preventable emergency room visits, and ultimately reduce all-cause readmission.</p> <p>In addition, our expert vascular surgeons and our community of vascular surgeons involved at the ACS contend that LEB surgery is a key leverage point for improving quality, and importantly, acknowledge there is variation in quality provided, thus facilitating mechanisms for improving care for the LEB population.</p> <p>If not measuring an outcome, provide evidence supporting this measure topic and grade the strength of the evidence <i>Summarize the evidence (including citations to source) supporting the focus of the measure as follows:</i></p> <ul style="list-style-type: none"> • Intermediate outcome - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit. • Process - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s). • Structure - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit. • Patient experience - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public. • Access - evidence that an association exists between access to a health service and the outcomes of, or experience with, care. • Efficiency- demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality. <p>Type of Evidence <i>Check all that apply</i></p> <p><input type="checkbox"/> Evidence-based guideline <input type="checkbox"/> Quantitative research studies</p>

² Citations can include, but are not limited to journal articles, reports, web pages (URLs).
 NQF Measure Submission Form, V3.0

	<input type="checkbox"/> Meta-analysis <input type="checkbox"/> Systematic synthesis of research <input type="checkbox"/> Qualitative research studies <input type="checkbox"/> Other (<i>Please describe</i>): <p>Overall Grade for Strength of the Evidence³ (<i>Use the USPSTF system, or if different, also describe how it relates to the USPSTF system</i>): Summary of Evidence (<i>provide guideline information below</i>): There is a large body of evidence that people can accomplish improvement to care of the general LEB population; bibliography is attached. Citations for Evidence: Attached</p>
21 (1c)	<p>Clinical Practice Guideline <i>Cite the guideline reference; quote the specific guideline recommendation related to the measure and the guideline author's assessment of the strength of the evidence; and summarize the rationale for using this guideline over others.</i></p> <p>Guideline Citation: N/A</p> <p>Specific guideline recommendation: N/A</p> <p>Guideline author's rating of strength of evidence (<i>If different from USPSTF, also describe it and how it relates to USPSTF</i>): N/A</p> <p>Rationale for using this guideline over others: N/A</p>
22 (1c)	<p>Controversy/Contradictory Evidence <i>Summarize any areas of controversy, contradictory evidence, or contradictory guidelines and provide citations.</i></p> <p>Summary: This is the first measure of its kind to be introduced to the NQF. The measure development team is not aware of any controversy/ contradictory evidence for this measure.</p> <p>Citations: N/A</p>
23 (1)	<p>Briefly describe how this measure (as specified) will facilitate significant gains in healthcare quality related to the specific priority goals and quality problems identified above: The measure evaluates patient safety outcomes (mortality and major complications) that are very important to the patient. Public reporting will drive quality improvement effort that result in reduced infection rates and improved outcomes for surgical patients.</p>
SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
Note: Testing and results should be summarized in this form. However, additional detail and reports may be submitted as supplemental information or provided as a web page URL. If a measure has not been tested, it is only potentially eligible for time-limited endorsement.	
24	Supplemental Testing Information: attached <input checked="" type="checkbox"/> OR Web page URL:
25 (2b)	<p>Reliability Testing</p> <p>Data/sample: Reliability testing will be done prior to measure implementation.</p> <p>Analytic Method: N/A</p> <p>Testing Results: N/A</p>

³The strength of the body of evidence for the specific measure focus should be systematically assessed and rated, e.g., USPSTF grading system www.ahrq.gov/clinic/uspstmeth.htm: A - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. B - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. C - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. D - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. I - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

<p>26 (2c)</p>	<p>Validity Testing</p> <p>Data/sample: The ACS NSQIP data elements are collected using electronic forms on the secure portion of NSQIP website. This data collection software has built-in checks for validity of data entries and consistencies between data fields. Additional data checking routines are run on the entire database before analyses are begun.</p> <p>Analytic Method:</p> <p>For simplicity, model validation was restricted to only the GLM. The hospitals in the study sample were randomly split into two sets: One for model development and the other for model validation. All patients from the chosen hospitals for each set were used in the data analysis. Using the development set, variables were selected using a stepwise selection procedure, and parameters were estimated. C-indices were calculated for both the development set and the validation set from the model using these parameter estimates. Similarity of the c-indices indicates validity of the model. Since both model development and model validation data files were small, this process was repeated ten times to assure consistent validation conclusions.</p> <p>Nagelkerke’s maximum (Nagelkerke et al., 2005), rescaled R-square (0.116), c-index (0.687), and Hosmer-Lemeshow test (Lemeshow et al., 1982), (p=0.469) indicate an adequately fit model with acceptable discrimination and predictive power. Repeated split-data analyses demonstrate similar variable selection with developmental data set c-indices between 0.684 and 0.716 and slightly lower c-indices for the test data sets, as expected, indicating model validity. The mean difference between c-indices for development and validation sets was 0.036 (SD 0.016).</p> <p>Testing Results: see above</p>
<p>27 (2d)</p>	<p>Measure Exclusions <i>Provide evidence to justify exclusion(s) and analysis of impact on measure results during testing.</i></p> <p>Summary of Evidence supporting exclusion(s): N/A</p> <p>Citations for Evidence: N/A</p> <p>Data/sample: N/A</p> <p>Analytic Method: N/A</p> <p>Testing Results: N/a</p>
<p>28 (2e)</p>	<p>Risk Adjustment Testing <i>Summarize the testing used to determine the need (or no need) for risk adjustment and the statistical performance of the risk adjustment method.</i></p> <p>Data/sample: The model was developed using 6,247 LEB patients in 152 hospitals with an overall unadjusted 30-day outcome (mortality or major complication) rate of 16.9%. The data used represents LEB procedures performed between July 1, 2004 and June 30, 2007. Patients undergoing a LEB were identified using Current Procedural Terminology (CPT) codes as per ACS NSQIP protocol (see list in question #5).</p> <p>Analytic Method: This risk-adjusted model predicts a composite outcome of mortality or major complications (within 30 days); model development was conducted in a manner that is tailored to and appropriate for a publicly reported outcome measures. Out analytic approach is similar to CMS’ previously developed acute myocardial infarction (AMI), heart failure (HF), and pneumonia (PN) 30-day risk standardized mortality measures.</p> <p>Clinical data such as lab values, included in ACS NSQIP data collection of 64 preoperative risk variables, were used to develop the model. A parsimonious set of predictors was identified using stepwise selection in a standard logistic regression model (generalized linear model or GLM) and expert clinical opinion.</p>

	<p>These predictors were then used to fit a GLM and a GLMM. Predictors were further validated by fitting a generalized estimating equations (GEE) model. Both GEE and GLMM account for clustering of patients within hospitals, but only GLMM has a formal random hospital effect. Model performance was assessed using Nagelkerke’s maximum rescaled R-square, c-index, and Hosmer-Lemeshow test. Validation of the final model was addressed by performing repeated split-data GLM analyses.</p> <p>Using the GLMM, parameter estimation was performed and a standardized outcome was determined for each hospital by calculating the ratio of predicted to expected mean outcomes (P/E ratio); the ratio of observed outcomes to expected outcomes (O/E ratio) was also calculated for both the GLM and GEE models. A 95% interval estimate was calculated for each P/E ratio, and a 95% confidence interval for each O/E ratio. The P/E and O/E ratios (and 95% intervals) were compared to further assess the use of GLMM as the final LEB model.</p> <p>Testing Results: Considering statistical and clinical significance, as well as data collection burden by hospitals, ten predictors were identified from the GLM model. The three multivariable prediction models (GLM, GLMM, and GEE) produced similar results for fixed effects when fitting models using the 10 predictors. Nagelkerke’s maximum rescaled R-square (0.116), c-index (0.687), and Hosmer-Lemeshow test (p=0.469) indicate an adequately fit model with acceptable discrimination and predictive power. Repeated split-data analyses demonstrate similar variable selection with developmental data set c-indices between 0.684 and 0.716 and slightly lower c-indices for the test data sets, as expected, indicating model validity. The mean difference between c-indices for development and validation sets was 0.036 (SD 0.016).</p> <p>The GLM and GEE model resulted in virtually identical hospital-level O/E ratios and hospital rankings. As expected, the extreme values of the standardized outcome using the GLM or the GEE model are compressed towards 1.0 in the GLMM model - particularly among hospitals with very low volume of patients.</p> <p>A parsimonious set of predictor variables was validated by finding similar results using GLM, GLMM, and GEE models. Use of GLMM was supported by the results of the O/E ratios for GLM/GEE and P/E ratios for GLMM, as the GLMM can provide estimates for low sample sizes, thereby including all providers in public reporting who would otherwise be excluded with GLM or GEE. Additionally, the GLMM model meets the standards set forth by the American Heart Association in 2006 for measures used for publicly reported outcomes, and parallels the methodology used in CMS’ AMI, HF, and PN mortality measures.</p> <p>► If outcome or resource use measure not risk adjusted, provide rationale: N/A</p>
<p>29 (2g)</p>	<p>Testing comparability of results when more than 1 data method is specified (<i>e.g., administrative claims or chart abstraction</i>)</p> <p>Data/sample: N/A</p> <p>Analytic Method: N/A</p> <p>Results: N/A</p>
<p>30 (2f)</p>	<p>Provide Measure Results from Testing or Current Use (<i>select one</i>)</p> <p>Data/sample: N/A</p> <p>Methods to identify statistically significant and practically/meaningfully differences in performance: N/A</p> <p>Results: N/A</p>
<p>31 (2h)</p>	<p>Identification of Disparities</p> <p>► If measure is stratified by factors related to disparities (i.e. race/ethnicity, primary language, gender, SES, health literacy), provide stratified results: N/A</p> <p>► If disparities have been reported/identified, but measure is not specified to detect disparities, provide</p>

	rationale: N/A
USABILITY	
32 (3)	<p>Current Use In development/testing If in use, how widely used (select one) ► If "other," please describe:</p> <p><input type="checkbox"/> Used in a public reporting initiative, name of initiative: Sample report attached <input type="checkbox"/> OR Web page URL:</p>
33 (3a)	<p>Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)</p> <p>Data/sample: N/A</p> <p>Methods: This project parallels the mortality and readmission measures that were previously submitted. In addition ACS has reported results of similar methods for several years.</p> <p>Results: N/A</p>
34 (3b, 3c)	<p>Relation to other NQF-endorsed™ measures ► Is this measure similar or related to measure(s) already endorsed by NQF (on the same topic or the same target population)? <i>Measures can be found at www.qualityforum.org under Core Documents.</i></p> <p><i>Check all that apply</i></p> <p><input type="checkbox"/> Have not looked at other NQF measures <input type="checkbox"/> Other measure(s) on same topic <input type="checkbox"/> Other measure(s) for same target population <input checked="" type="checkbox"/> No similar or related measures</p> <p>Name of similar or related NQF-endorsed™ measure(s): No similar NQF endorsed measure.</p> <p>Are the measure specifications harmonized with existing NQF-endorsed™ measures? Yes, fully harmonized</p> <p>► If not fully harmonized, provide rationale: According to the NQF list of endorsed measures and the list of candidate measures dated May 15, 2008, there are no risk adjusted outcome measures for patients undergoing a lower extremity bypass procedure.</p> <p>Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:</p>
FEASIBILITY	
35 (4a)	<p>How are the required data elements generated? <i>Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data elements are generated concurrent with and as a byproduct of care processes during care delivery (e.g., blood pressure or other assessment recorded by personnel conducting the assessment)</p> <p><input type="checkbox"/> Data elements are generated from a patient survey (e.g., CAHPS)</p> <p><input type="checkbox"/> Data elements are generated through coding performed by someone other than the person who obtained the original information (e.g., DRG or ICD-9 coding on claims)</p> <p><input checked="" type="checkbox"/> Other, Please describe: medical chart abstraction during admission, post-discharge information collected using patient follow-up.</p>
36 (4b)	<p>Electronic Sources All data elements</p> <p>► If all data elements are not in electronic sources, specify the near-term path to electronic collection by most providers: Using data collection tool to capture required data elements from medical chart</p> <p>► Specify the data elements for the electronic health record: see technical report</p>
37 (4c)	<p>Do the specified exclusions require additional data sources beyond what is required for the other specifications? No</p> <p>► If yes, provide justification: N/A</p>
38	Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure:

(4d)	<p>Data inaccuracy is typically an issue when collecting data. However, using a well-designed data collection tool that uses a minimum number of necessary data elements would minimize the data collection burden and thus minimize the potential for errors.</p> <p>One universal potential unintended consequence of the measuring and publicly reporting important quality measures, such mortality and complications is the notion that hospitals and/or clinicians would begin to turn away very sick patients in the interest of protecting their scores. The risk adjustment mitigates this concern.</p> <p>Describe how could these potential problems be audited:</p> <p>Did you audit for these potential problems during testing? Yes If yes, provide results: The ACS NSQIP data elements are collected using electronic forms on the secured portion of the NSQIP website. This data data collection software has build-it checks for validity of data entries and consistencies between data fields. The reliability of data abstraction is determined by re-abstraction of a sample of cases by nurses from the company responsible for the development and maintenance of the data collection software.</p>
39 (4e)	<p>Testing feasibility Describe what have you learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: The feasibility of hospitals being able to participate and collect data elements has already been demonstrated by the 200-250 hospitals currently participating in ACS NSQIP.</p>
CONTACT INFORMATION	
40	<p>Web Page URL for Measure Information Describe where users (implementers) should go for more details on specifications of measures, or assistance in implementing the measure. Web page URL: qualitynet.org OR acsnsqip.org</p>
41	<p>Measure Intellectual Property Agreement Owner Point of Contact First Name: MI: Last Name: Credentials (MD, MPH, etc.): Organization: Street Address: City: State: ZIP: Email: Telephone: ext:</p>
42	<p>Measure Submission Point of Contact If different than IP Owner Contact First Name: Lein MI: F Last Name: Han Credentials (MD, MPH, etc.): PhD Organization: Centers for Medicare & Medicaid Services (CMS) Street Address: 7500 Security Blvd City: Baltimore State: MD ZIP: 21244-9045 Email: lein.han@cms.hhs.gov Telephone: 410-786-0205 ext:</p>
43	<p>Measure Developer Point of Contact If different than IP Owner Contact First Name: MI: Last Name: Credentials (MD, MPH, etc.): Organization: Street Address: City: State: ZIP: Email: Telephone: ext:</p>
44	<p>Measure Steward Point of Contact If different than IP Owner Contact <i>Identifies the organization that will take responsibility for updating the measure and assuring it is consistent with the scientific evidence and current coding schema; the steward of the measure may be different than the developer.</i> First Name: Lein MI: F Last Name: Han Credentials (MD, MPH, etc.): PhD Organization: Centers for Medicare & Medicaid Services (CMS) Street Address: 7500 Security Blvd City: Baltimore State: MD ZIP: 21244-9045 Email: lein.han@cms.hhs.gov Telephone: 410-786-0205 ext</p>

ADDITIONAL INFORMATION	
45	<p>Workgroup/Expert Panel involved in measure development Workgroup/panel used</p> <ul style="list-style-type: none"> ▶ If workgroup used, describe the members' role in measure development: ▶ Provide a list of workgroup/panel members' names and organizations: <p>Colorado Foundation for Medical Care</p> <p>Dima Turkmani, MBA, MPH Project Director</p> <p>Maureen O'Brien, PhD Senior Scientist</p> <p>Beth Stevens, MS Biostatistician</p> <p>Mary Hajner, BA Project Coordinator</p> <p>American College of Surgeons</p> <p>Clifford Ko, MD, MS, MSHS Director of Research and Optimal Patient Care Surgeon, UCLA Medical Center</p> <p>Karen Richards, BS Administrative Director</p> <p>Sameera Ali Grants Manager</p> <p>Colorado Health Outcomes Program, University of Colorado Denver</p> <p>Karl E. Hammermeister, MD Professor of Medicine Colorado Health Outcomes Program School of Medicine University of Colorado</p> <p>William Henderson, PhD, MPH Professor (Biostatistics) Colorado Health Outcomes Program School of Public Health University of Colorado</p> <p>Sung-joon Min, PhD Biostatistician Assistant Professor of Medicine (Division of Health Care Policy and Research) University of Colorado Denver</p> <p>Patrick Hosokawa, MS Biostatistician</p> <p>Washington University, St. Louis</p>

	<p>Bruce Hall, MD, PhD, MBA Department of Surgery, John Cochran Veterans Affairs Medical Center, St Louis, MO. Department of Surgery, School of Medicine; Washington University in St Louis, St Louis, MO. Olin Business School; and Center for Health Policy, Washington University in St Louis, St Louis, MO.</p> <p>The primary workgroup has consulted with the following experts:</p> <p>Sharon-Lise Normand, PhD Biostatistician and Professor of Health Care Policy (Biostatistics) Harvard Medical School</p> <p>Julie Freischlag, MD Professor of Surgery The Johns Hopkins University School of Medicine</p> <p>Anton Sidawy, MD President of the American Society for Clinical Vascular Surgery Chief of Surgery at the Veterans Affairs Medical Center in Washington DC Professor of Surgery at both Georgetown and George Washington University Schools of Medicine</p>
46	<p>Measure Developer/Steward Updates and Ongoing Maintenance Year the measure was first released: N/A Month and Year of most recent revision: N/A What is the frequency for review/update of this measure? N/A When is the next scheduled review/update for this measure? N/A</p>
47	<p>Copyright statement/disclaimers:</p>
48	<p>Additional Information: LEB measure methodology report is attached. References include:</p> <p>Laird NM, Ware JH. Random-effects models for longitudinal data. <i>Biometrics</i> 1982;38(4):963-74.</p> <p>Landis JR, Koch GG. The measurement of observer agreement for categorical data. <i>Biometrics</i> 1977;33(1):159-74.</p> <p>Lemeshow S, Hosmer DW, Jr. A review of goodness of fit statistics for use in the development of logistic regression models. <i>Am J Epidemiol</i> 1982;115(1):92-106.</p> <p>Molenberghs G, Verbeke G. <i>Models for Discrete Longitudinal Data</i>. New York: Springer; 2005.</p> <p>Nagelkerke N, Smits J, Ie CS, van HH. Testing goodness-of-fit of the logistic regression model in case-control studies using sample reweighting. <i>Stat Med</i> 2005 Jan 15;24(1):121-30</p> <p>Raudenbush SW, Bryk AS. <i>Hierarchical Linear Models: Applications and Data Analysis Methods</i>. 2nd ed. Thousand Oaks, CA: Sage Publications; 2002</p> <p>Snijders TA, Bosker RJ. <i>Multilevel Analysis: An introduction to basic and advanced multilevel modeling</i>. London: Sage Publications; 2000</p>
49	<p>I have checked that the submission is complete and any blank fields indicate that no information is provided. <input checked="" type="checkbox"/></p>
50	<p>Date of Submission (MM/DD/YY): 11-21-08</p>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.0

August 2008

The measure information you submit will be shared with NQF’s Steering Committees and Technical Advisory Panels to evaluate measures against the NQF criteria of importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Four conditions (as indicated below) must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, importance to measure and report, in order to be evaluated against the remaining criteria. References to the specific measure evaluation criteria are provided in parentheses following the item numbers. Please refer to the *Measure Evaluation Criteria* for more information at www.qualityforum.org under Core Documents. Additional guidance is being developed and when available will be posted on the NQF website.

Use the tab or arrow (↓→) keys to move the cursor to the next field (or back ←↑). There are three types of response fields:

- drop-down menus - select one response;
- check boxes - check as many as apply; and
- text fields - you can copy and paste text into these fields or enter text; these fields are not limited in size, but in most cases, we ask that you summarize the requested information.

Please note that URL hyperlinks do not work in the form; you will need to type them into your web browser.

Be sure to answer all questions. Fields that are left blank will be interpreted as no or none. Information must be provided in this form. Attachments are not allowed except when specifically requested or to provide additional detail or source documents for information that is summarized in this form. If you have important information that is not addressed by the questions, they can be entered into item #48 near the end of the form.

For questions about this form, please contact the NQF Project Director listed in the corresponding call for measures.

CONDITIONS FOR CONSIDERATION BY NQF	
	<i>Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards.</i>
A (A)	<i>Public domain or Intellectual Property Agreement signed: Public domain - IP agreement not required (If no, do not submit) Template for the Intellectual Property Agreement is available at www.qualityforum.org under Core Documents.</i>
B (B)	<i>Measure steward/maintenance: Is there an identified responsible entity and process to maintain and update the measure on a schedule commensurate with clinical innovation, but at least every 3 years? Yes, information provided in contact section (If no, do not submit)</i>
C (C)	<i>Intended use: Does the intended use of the measure include BOTH public reporting AND quality improvement? Yes (If no, do not submit)</i>
D (D)	<i>Fully developed and tested: Is the measure fully developed AND tested? No, testing will be completed within 24 months (If not tested and no plans for testing within 24 months, do not submit)</i>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.0

August 2008

	<i>(for NQF staff use) NQF Review #: HOE-009-08</i> <i>NQF Project: Hospital Outcomes and Efficiency</i>
MEASURE SPECIFICATIONS & DESCRIPTIVE INFORMATION	
1	<i>Information current as of (date- MM/DD/YY): 11-21-08</i>
2	<i>Title of Measure: 30-day all-cause risk-standardized percutaneous coronary intervention (PCI) mortality rate for patients without ST segment elevation myocardial infarction (STEMI) and without cardiogenic shock</i>
3	Brief description of measure ¹ : Hospital-specific 30-day all-cause risk-standardized mortality rate following Percutaneous Coronary Intervention (PCI) among patients aged 18 years or older without ST segment elevation myocardial infarction (STEMI) and without cardiogenic shock at the time of procedure.
4 (2a)	<p>Numerator Statement: Note: This outcome measure does not have a traditional numerator and denominator like a core process measure (e.g., percentage of adult patients with diabetes aged 18-75 years receiving one or more hemoglobin A1c tests per year); thus, we use this field to define our statistically-adjusted rate outcome measure.</p> <p>We use hierarchical logistic regression modeling to calculate a hospital-specific 30-day risk-standardized mortality rate (RSMR). This rate is calculated as the ratio of “predicted” to “expected” deaths, multiplied by the national unadjusted mortality rate. For each hospital, the “numerator” of the ratio component of the RSMR is the predicted number of deaths within 30 days given the hospital’s performance with its observed case mix, and the “denominator” is the expected number of deaths given the hospital’s case mix. By convention, we use the term “predicted” here to describe the numerator result, which is calculated using the hospital-specific intercept term. We use “expected” for the denominator, which is calculated using the average intercept term.</p> <p>More specifically, the expected number of deaths for each hospital is estimated using its patient mix and the average hospital-specific intercept. The predicted number of deaths for each hospital is estimated given the same patient mix but the hospital-specific intercept. Operationally, the expected number of deaths for each hospital is obtained by regressing the risk factors (see # 8) on the death using all hospitals in our sample, applying the subsequent estimated regression coefficients to the patient characteristics observed in the hospital, adding the average of the hospital-specific intercepts, transforming, and then summing over all patients in the hospital to get a value. This is a form of indirect standardization. The predicted hospital outcome is the number of deaths in the “specific” hospital estimated given its performance and case mix. Operationally, this is accomplished by estimating a hospital-specific intercept that represented baseline mortality risk within the hospital, applying the estimated regression coefficients to the patient characteristics in the hospital, transforming, and then summing over all patients in the hospital to get a value. To assess hospital performance in any given year, we re-estimate the model coefficients using that year’s data. (Please see the attached methodology report for details of the statistical methodology.)</p> <p>Time Window: This measure was developed with 24 months of data. The time period for public reporting has not been determined.</p> <p>Numerator Details (Definitions, codes with description):</p>
5 (2a)	<p>Denominator Statement: For “Denominator Statement,” please see #4 above. Instead, we are using this field to define our patient cohort.</p> <p>Outcome measure cohort definition: PCI procedures for patients at least 18 years of age, without STEMI</p>

¹ Example of measure description: Percentage of adult patients with diabetes aged 18-75 years receiving one or more A1c test(s) per year.
NQF Measure Submission Form, V3.0

	<p>and without cardiogenic shock at the time of procedure.</p> <p>Time Window: This measure was developed with 24 months of data. The time period for public reporting has not been determined.</p> <p>Denominator Details (Definitions, codes with description): See above. We are using this field to specify the codes that define the PCI patient cohort.</p> <p>In the CathPCI Registry, admissions with PCI are identified by field 614 (PCI=yes); STEMI and shock are defined as follows: (1) Symptoms present on admission = ACS:STEMI (field 550 = 6) with Time Period Symptom Onset to Admission within 24 hours (field 560 = 1,2,3) or Acute PCI = Yes (field 812 = 2,3,4); OR (2) Cardiogenic shock = Yes (field 520=1).</p> <p>All patients who do not meet any of the above criteria are patients with no STEMI within 24 hours of arrival to the hospital and no cardiogenic shock prior to the PCI. These patients are included in the without STEMI and without shock cohort.</p>
<p>6 (2a, 2d)</p>	<p>Denominator Exclusions: Note: We are using this field to define exclusions to the patient cohort.</p> <p>(1) PCIs that follow a prior PCI in the same admission or occur during a transfer-in admission (PCI to PCI). We define an episode of care as starting on the day of the PCI during the first admission regardless of whether additional procedures are performed at the same hospital or at a different hospital after transfer. Thus, in the period of evaluation after the index procedure we do not begin a new period of evaluation after a second PCI during the same episode of care. If the patient is discharged to a non-acute care facility and has a second PCI within 30-days, that PCI is eligible as a new index PCI (except as noted in 3 below).</p> <p>(2) PCIs in patients with missing vital status (inability to link patient information to appropriate death index). In actual practice, with the identifiers that will be collected as part of the database we anticipate that missing data will be rare.</p> <p>(3) PCIs which would lead to duplicate attribution of 30-day deaths. The 30-day outcome period for patients with more than one PCI may overlap. In order to avoid attributing the same death to more than one PCI (i.e. double counting a single patient death), later PCI procedures within 30 days of the death are excluded.</p> <p>(4) PCIs for patients with more than 10 days between date of admission and date of PCI. Patients who have a PCI after many days of hospitalization are rare and represent a distinct population that likely has risk factors related to the hospitalization that are not well quantified in the registry. It seemed clinically sensible to exclude these patients.</p> <p>Denominator Exclusion Details (Definitions, codes with description): See above. We are deriving the corresponding codes based on the data for exclusion.</p>
<p>7 (2a, 2h)</p>	<p>Stratification Do the measure specifications require the results to be stratified? No ▶ If “other” describe:</p> <p>Identification of stratification variable(s):</p> <p>Stratification Details (Definitions, codes with description): This measure was designed to be reported along with 30-day all-cause risk-standardized percutaneous coronary intervention (PCI) mortality rate for patients with ST segment elevation myocardial infarction (STEMI) or cardiogenic shock</p>
<p>8 (2a, 2e)</p>	<p>Risk Adjustment Does the measure require risk adjustment to account for differences in patient severity before the onset of care? Yes ▶ If yes, Statistical Risk Model, see Variables ▶ Is there a separate proprietary owner of the risk model? No</p> <p>Identify Risk Adjustment Variables:</p>

	<p>Age (10 year increments) Body Mass Index (5 kg/m² increments) Heart Failure - Previous History Cerebrovascular disease Peripheral Vascular Disease Chronic Lung disease Diabetes/Control 0=No Diabetes 1=Non-Insulin Diabetes 2=Insulin Diabetes Glomerular Filtration Rate (GFR) (derived) 0=Not measured 1="GFR<30" 2="30≤GFR<60" 3="60≤GFR<90" 4="GFR≥90" Previous PCI Heart Failure - Current Status NYHA: Class IV Symptom Onset No MI on admission MI within 24 hours of admission MI > 24 hours after admission Ejection Fraction Percent (EF) 1=Not measured 2="EF<30" 3="30≤ EF<45" 4="EF≥45" PCI status 1=Elective 2=Urgent 3=Emergency or 4=Salvage Highest Risk Lesion - coronary artery segment category 1=proximal RCA/mid LAD/proximal Cx 2=proximal LAD 3=Left Main Highest Risk Lesion: Society for Cardiovascular Angiography and Interventions (SCAI) class 1 class 2 or 3 class 4</p> <p>For more details, please see the attached methodology report.</p> <p>Detailed risk model: attached <input checked="" type="checkbox"/> OR Web page URL:</p>
<p>9 (2a)</p>	<p>Type of Score: Rate/proportion Calculation Algorithm: attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p>Interpretation of Score (<i>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</i>) Other ► If "Other", please describe: For each hospital, we calculate a 30-day all-cause risk-standardized mortality rate (RSMR) and the 95% interval estimate for the RSMR, which expresses the level of uncertainty around the point estimate. The RSMR with its interval estimate can be used to classify hospital performance (e.g., higher than expected, as expected, or lower than expected).</p>
<p>10 (2a.)</p>	<p>Identify the required data elements(e.g., primary diagnosis, lab values, vital signs): Data dictionary/code table attached <input type="checkbox"/> OR Web page URL: http://www.ncdr.com/WebNCDR/ELEMENTS.ASPX</p>

4a, 4b)	<p>Data Quality (2a) <i>Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data are captured from an authoritative/accurate source (e.g., lab values from laboratory personnel)</p> <p><input checked="" type="checkbox"/> Data are coded using recognized data standards</p> <p><input type="checkbox"/> Method of capturing data electronically fits the workflow of the authoritative source</p> <p><input type="checkbox"/> Data are available in EHRs</p> <p><input checked="" type="checkbox"/> Data are auditable</p>
11	<p>Data Source and Data Collection Methods <i>Identifies the data source(s) necessary to implement the measure specifications. Check all that apply</i></p> <p>(2a, 4b) <input type="checkbox"/> Electronic Health/Medical Record <input type="checkbox"/> Paper Medical Record</p> <p><input type="checkbox"/> Electronic Clinical Database, Name: <input type="checkbox"/> Standardized clinical instrument, Name:</p> <p><input checked="" type="checkbox"/> Electronic Clinical Registry, Name: National Cardiovascular Data Registry, CathPCI Registry; required data could alternatively be collected through other non-registry mechanisms <input type="checkbox"/> Standardized patient survey, Name:</p> <p><input type="checkbox"/> Electronic Claims <input type="checkbox"/> Standardized clinician survey, Name:</p> <p><input type="checkbox"/> Electronic Pharmacy data <input checked="" type="checkbox"/> Other, Describe: Death Index</p> <p><input type="checkbox"/> Electronic Lab data Instrument/survey attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p><input type="checkbox"/> Electronic source - other, Describe:</p>
12	<p>Sampling <i>If measure is based on a sample, provide instructions and guidance on sample size.</i></p> <p>(2a) Minimum sample size: Data from all hospitals and all PCI would be included in the process of reestimating model variables. For public reporting, minimum sample size has not been determined.</p> <p>Instructions: N/A</p>
13	<p>Type of Measure: Outcome ▶ If “Other”, please describe:</p> <p>(2a) ▶ If part of a composite or paired with another measure, please identify composite or paired measure This measure is being submitted along with: 30-day all-cause risk-standardized percutaneous coronary intervention (PCI) mortality rate for patients with ST segment elevation myocardial infarction (STEMI) or cardiogenic shock.</p>
14	<p>Unit of Measurement/Analysis <i>(Who or what is being measured) Check all that apply.</i></p> <p>(2a) <input type="checkbox"/> Can be measured at all levels <input type="checkbox"/> Integrated delivery system</p> <p><input type="checkbox"/> Individual clinician (e.g., physician, nurse) <input type="checkbox"/> Health plan</p> <p><input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice) <input type="checkbox"/> Community/Population</p> <p><input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home) <input type="checkbox"/> Other <i>(Please describe):</i></p>
15	<p>Applicable Care Settings <i>Check all that apply</i></p> <p>(2a) <input type="checkbox"/> Can be used in all healthcare settings <input type="checkbox"/> Hospice</p> <p><input type="checkbox"/> Ambulatory Care (office/clinic) <input checked="" type="checkbox"/> Hospital</p> <p><input type="checkbox"/> Behavioral Healthcare <input type="checkbox"/> Long term acute care hospital</p> <p><input type="checkbox"/> Community Healthcare <input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)</p> <p><input type="checkbox"/> Dialysis Facility <input type="checkbox"/> Prescription Drug Plan</p> <p><input type="checkbox"/> Emergency Department <input type="checkbox"/> Rehabilitation Facility</p> <p><input type="checkbox"/> EMS emergency medical services <input type="checkbox"/> Substance Use Treatment Program/Center</p> <p><input type="checkbox"/> Health Plan <input type="checkbox"/> Other <i>(Please describe):</i></p> <p><input type="checkbox"/> Home Health</p>
IMPORTANCE TO MEASURE AND REPORT	
Note: This is a threshold criterion. If a measure is not judged to be sufficiently important to measure and report, it will not be evaluated against the remaining criteria.	
16	<p>Addresses a Specific National Priority Partners Goal <i>Enter the numbers of the specific goals related to this measure (see list of goals on last page):</i> 3.3, 3.4</p>
17	<p>If not related to NPP goal, identify high impact aspect of healthcare (select one)</p>

(1a)	<p>Summary of Evidence:</p> <p>Citations² for Evidence:</p>
18 (1b)	<p>Opportunity for Improvement <i>Provide evidence that demonstrates considerable variation, or overall poor performance, across providers.</i></p> <p>Summary of Evidence: PCI is one of the most commonly performed cardiac procedures in the United States. In 2005, an estimated 1,265,000 PCI procedures were performed in the United States (Rosamond Flegal et al. 2008). From 1987-2003, the number of procedures increased 326% (Thom, Haase et al. 2006). Inpatient mortality is the indicator that has been most widely used to evaluate the quality of cardiac procedures and is arguably the most important adverse outcome measure. The ACC summarized the experience of the NCDR CathPCI Registry from 1998-2000 and found that in-hospital mortality occurred in 1,422 of 100,253 PCI procedures (1.4%) (Shaw, Anderson et al. 2002). In the present era, mortality rates for PCI in large series from experienced operators ranged from 0.5 to 1.7 percent (Carrozza, Cutlip et al. 2008). Prior studies have demonstrated significant variability in in-hospital PCI mortality across age groups, gender, geographic regions, socioeconomic status, and by hospital volume (Mukherjee, Wainess et al. 2005). Although 12 states already report PCI outcomes, to date there has not been a unified national effort to publicly report PCI mortality.</p> <p>Citations for Evidence:</p> <p>Carrozza J, Cutlip D, Levin T. (2008). Periprocedural complications of percutaneous coronary intervention. UpToDate. B. Rose. Waltham, MA.</p> <p>Mukherjee D, Wainess RM, et al. (2005). "Variation in outcomes after percutaneous coronary intervention in the United States and predictors of periprocedural mortality." <i>Cardiology</i> 103(3): 143-7.</p> <p>Rosamond W, Flegal K, Furie K, Go A, Greenlund K, Haase N, Hailpern SM, Ho M, Howard V, Kissela B, Kittner S, Lloyd-Jones D, McDermott M, Meigs J, Moy C, Nichol G, O'Donnell C, Roger V, Sorlie P, Steinberger J, Thom T, Wilson M, Hong Y. Heart Disease and Stroke Statistics_2008 Update: A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee and for the American Heart Association Statistics Committee and Stroke Statistics Subcommittee <i>Circulation</i> 2008;117:e25-e146; originally published online Dec 17, 2007; DOI: 10.1161/CIRCULATIONAHA.107.187998.</p> <p>Shaw RE, Anderson HV, et al. (2002). "Development of a risk adjustment mortality model using the American College of Cardiology-National Cardiovascular Data Registry (ACC-NCDR) experience: 1998-2000." <i>J Am Coll Cardiol</i> 39(7): 1104-12.</p> <p>Thom T, Haase N, et al. (2006). "Heart disease and stroke statistics--2006 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee." <i>Circulation</i> 113(6): e85-151.</p>
19 (1b)	<p>Disparities <i>Provide evidence that demonstrates disparity in care/outcomes related to the measure focus among populations.</i></p> <p>Summary of Evidence: We have not examined health disparities associated with this measure. This measure could be used to assess differences in performance among hospitals that care for different types of populations (e.g., those that serve primarily minority populations versus others).</p> <p>Citations for evidence: N/A</p>
20 (1c)	<p>If measuring an Outcome Describe relevance to the national health goal/priority, condition, population, and/or care being addressed: This measure will describe hospital-level mortality rates following PCI with the overriding goal to reduce preventable and premature mortality rates to best-in-class (NPP 3.3) and 30-day mortality rates following hospitalization for relevant conditions to best-in-class (NPP 3.4).</p> <p>If not measuring an outcome, provide evidence supporting this measure topic and grade the strength</p>

² Citations can include, but are not limited to journal articles, reports, web pages (URLs).
NQF Measure Submission Form, V3.0

of the evidence
Summarize the evidence (including citations to source) supporting the focus of the measure as follows:

- **Intermediate outcome** - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
- **Process** - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
- **Structure** - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
- **Patient experience** - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
- **Access** - evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
- **Efficiency**- demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

Type of Evidence *Check all that apply*

<input type="checkbox"/> Evidence-based guideline	<input type="checkbox"/> Quantitative research studies
<input type="checkbox"/> Meta-analysis	<input type="checkbox"/> Qualitative research studies
<input type="checkbox"/> Systematic synthesis of research	<input type="checkbox"/> Other (Please describe):

Overall Grade for Strength of the Evidence³ (Use the USPSTF system, or if different, also describe how it relates to the USPSTF system): **N/A**

Summary of Evidence (provide guideline information below):
Evidence that the outcome measure has been influenced by one or more clinical interventions:

Numerous studies have demonstrated the efficacy of interventions designed to improve patient outcomes following PCI. These include pharmacologic interventions such as the use of glycoprotein 2b/3a inhibitors, direct thrombin inhibitors, and pre-procedural clopidogrel, as well as advances in device technology such as use of stents (and more recently drug eluting stents), thrombectomy for acute lesions with high thrombus burden, and distal embolic protection for PCI of degenerated saphenous vein grafts. Of note, the majority of these interventions have been shown to reduce endpoints other than mortality, most commonly rates of periprocedural MI, major bleeding, and target vessel revascularization for in-stent restenosis. Although few individual interventions have been shown to reduce mortality, they may collectively exert a favorable impact on hospital mortality rates following PCI when implemented in a coordinated fashion.

There is a growing body of evidence that quality improvement efforts can improve outcomes of PCI patients, including survival. Rihal and colleagues examined patient outcomes before and after initiation of a program of continuous quality improvement (CQI) and found a significantly lower in-hospital mortality following PCI despite significant increases in the risk profile of PCI patients. Similar improvements were identified in studies of CQI by Brush et al and Moscucci et al, and improvements in survival were associated with greater adherence to evidence based practices including preprocedural clopidogrel, use of glycoprotein 2b/3a inhibitors, and volume of iodinated contrast. The observational nature of these studies precludes drawing definitive conclusions, but they strongly suggest a mechanism by which public reporting of hospital PCI outcomes could promote improvements in the care of PCI patients.

Citations for Evidence: **Brush JE, Balakrishnan SA, Brough J, Hartman C, Hines G, Liverman DP, Parker JP,**

³The strength of the body of evidence for the specific measure focus should be systematically assessed and rated, e.g., USPSTF grading system www.ahrq.gov/clinic/uspstmeth.htm: A - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. B - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. C - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. D - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. I - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

	<p>Rich J, Tindall N. (2006). "Implementation of a continuous quality improvement program for percutaneous coronary intervention and cardiac surgery at a large community hospital." <i>Am Heart J</i> 152 (2):379-85 16875926 (P,S,E,B).</p> <p>Moscucci M, Kline Rogers E, Montoye C, Smith DE, Share D, O'Donnell M, Maxwell-Eward A, Meengs WL, De Franco AC, Patel K, McNamara R, McGinnity JG, Jani SM, Khanal S, Eagle KA. (2006). "Association of a Continuous Quality Improvement Initiative With Practice and Outcome Variations of Contemporary Percutaneous Coronary Interventions." <i>Circulation</i>. 113:814-822.</p> <p>Rihal C, Kamath C, Holmes D, et al. (2006). "Economic and clinical outcomes of a physician-led continuous quality improvement intervention in the delivery of percutaneous coronary intervention." <i>Am J Manag Care</i> 12:445-452.</p>
<p>21 (1c)</p>	<p>Clinical Practice Guideline <i>Cite the guideline reference; quote the specific guideline recommendation related to the measure and the guideline author's assessment of the strength of the evidence; and summarize the rationale for using this guideline over others.</i></p> <p>Guideline Citation: N/A</p> <p>Specific guideline recommendation: N/A</p> <p>Guideline author's rating of strength of evidence (If different from USPSTF, also describe it and how it relates to USPSTF): N/A</p> <p>Rationale for using this guideline over others: N/A</p>
<p>22 (1c)</p>	<p>Controversy/Contradictory Evidence <i>Summarize any areas of controversy, contradictory evidence, or contradictory guidelines and provide citations.</i></p> <p>Summary: There are a few points of potential controversy that deserve comment:</p> <p>1) This model is designed for use in national public reporting and is aligned with the American Heart Association (AHA) published standards for publicly reported outcomes measures (Krumholz et al. 2006). The model, however, was developed from a subset of the entire population of PCI patients, namely fee-for-service Medicare patients undergoing PIC at facilities that participate in the NCDR CathPCI Registry. Furthermore, patients' vital status at 30 days was determined by linking to administrative data using a probabilistic match because our derivation and validation samples lacked unique identifiers to merge with a national death index.</p> <p>For public reporting, the parameters would be re-estimated using the national data. In addition, direct identifiers would be used to link clinical data and determine vital status. Further, adequate mechanisms would need to be implemented to ensure data quality (such as monitoring data for variances in case mix [e.g., unexpectedly high proportion of salvage PCI or cardiogenic shock], chart audits, and possibly adjudicating cases that are vulnerable to systematic misclassification). There is no reason to believe that these changes will significantly change the performance of the model.</p> <p>2) We chose 30-day mortality as our period of assessment. In contrast, prior efforts to create models for risk adjusting PCI outcomes have used hospitals' self-reported in-hospital mortality. Advantages of a 30-day mortality outcome include providing a standardized period of assessment, potentially more accurate assessments of vital status, and providing a more complete picture of outcomes following PCI. The main disadvantage of this approach is that it requires linking clinical data to a separate data source to determine 30-day vital status. In order to inform this decision, we determined whether in-hospital mortality rates were comparable to 30-day mortality rates. We found that the median absolute difference in unadjusted mortality rates was 0.5%, but that 8% of hospitals had a greater than 2% absolute difference. Furthermore, when we compared hospitals' decile ranking using these two endpoints, a quarter of hospitals changed more than one decile. Based on this evidence, we determined that in-hospital mortality may not be an adequate surrogate for 30-day mortality following PCI.</p> <p>3) We propose models that stratify the population of patients undergoing PCI into two distinct cohorts:</p>

	<p>patients with STEMI or cardiogenic shock, and patients without STEMI and without shock. This approach reflects the fact that among patients undergoing PCI, the risk of mortality differs considerably depending on the clinical context in which it is performed. The mortality of PCI patients with an evolving STEMI is substantially higher than that of outpatients undergoing elective procedures. Furthermore, many hospitals (e.g., primary PCI centers) can only perform PCI on patients undergoing non-elective procedures. Stratifying the population into these cohorts was felt to provide more fair and accurate comparisons of the outcomes of patients treated at different types of hospitals. This strategy has previously been implemented by the Massachusetts program for publicly reporting of mortality following PCI (www.massdac.org/pic/index.htm). The state of New York reports outcomes for both the combined cohort as well as a stratified cohort (www.health.state.ny.us/statistics/diseases/cardiovascular).</p> <p>4) We did not consider as candidate variables those that we would not want to adjust for in a quality measure, such as potential complications, certain patient demographics (e.g., race, socioeconomic status), and patients' admission path (e.g., admitted from, or discharged to, a skilled nursing facility [SNF]). These characteristics may be associated with mortality and thus could increase the model performance to predict patient mortality. However, these variables may be related to quality or supply factors that should not be included in an adjustment that seeks to control for patient clinical characteristics while illuminating important quality differences.</p> <p>5) Studies suggest that public reporting of the outcomes of cardiovascular procedures may have unintended consequences. Moscucci and colleagues compared the characteristics and outcomes of patients undergoing PCI in states with (New York) and without (Michigan) public reporting and found that patients undergoing PCI in New York were substantially lower risk than PCI patients in Michigan. Determining the underlying causes and appropriateness of these differences is impossible, but there is concern that physicians in states that publicly report PCI outcomes would either refer high risk cases to states without public reporting or avoid such cases altogether. Implementing a national measure of PCI outcomes would avoid the former problem in that public reporting would be consistent across states. Nevertheless, the proposed measure will require close attention to the possibility that high risk patients are not receiving PCI when clinically indicated. The proposed measure is, however, complementary to the previously approved measures for 30-day mortality of AMI and heart failure patients in that inappropriate avoidance of high risk PCI cases may have a detrimental effect on hospitals' performance on these other measures of cardiovascular outcomes.</p> <p>Citations: Krumholz HM, Brindis RG, et al. (2006). "Standards for statistical models used for public reporting of health outcomes: an American Heart Association Scientific Statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation." <i>Circulation</i> 113(3): 456-62.</p>
23 (1)	<p>Briefly describe how this measure (as specified) will facilitate significant gains in healthcare quality related to the specific priority goals and quality problems identified above: Public reporting will drive internal hospital quality improvement efforts to achieve mortality rates consistent with or better than the best performing hospitals in the country.</p>
SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
<p>Note: Testing and results should be summarized in this form. However, additional detail and reports may be submitted as supplemental information or provided as a web page URL. If a measure has not been tested, it is only potentially eligible for time-limited endorsement.</p>	
24	<p>Supplemental Testing Information: attached <input checked="" type="checkbox"/> OR Web page URL:</p>
25 (2b)	<p>Reliability Testing</p> <p>Data/sample: N/A</p> <p>Analytic Method: N/A</p> <p>Testing Results: Reliability testing using data from all PCIs without STEMI and without Shock will be</p>

	performed prior to measure implementation.
26 (2c)	<p>Validity Testing</p> <p>Data/sample: We developed this model in a limited cohort (see section 22 above).</p> <p>We developed a model in a cohort of NCDR patients ≥ 65 years old who had matching information in Medicare administrative data that allowed us to link to the outcome (mortality within 30 days). In the Medicare dataset, admissions with PCI are identified by International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) procedure codes, shown here:</p> <ul style="list-style-type: none"> 00.66 Percutaneous transluminal coronary angioplasty or coronary atherectomy 36.01 Single vessel PTCA or coronary atherectomy 36.02 Percutaneous transluminal coronary angioplasty or coronary atherectomy with mention of thrombolytic agent 36.05 Multiple vessel PTCA or coronary atherectomy 36.06 Insertion of non-drug-eluting coronary artery stent(s) 36.07 Insertion of drug-eluting coronary artery stent(s) <p>Analytic Method: Because patient identifiers are not currently available, we used a probabilistic match to link the two datasets using hospital Medicare Provider Number (MPN), patient age, gender, date of admission, and date of discharge. We matched 65% of the PCI cohort after excluding non-unique records. Matched and unmatched patients had similar clinical characteristics.</p> <p>For more details, please see the attached methodology report.</p> <p>Overview: The evidence supporting the measure can be found in the Methodology report accompanying this submission. In brief, a risk adjustment model was derived using all matched admissions in 2006 (“development sample”). The performance of the models was validated using a similar cohort of patients who underwent PCI in 2005 (“validation sample”). For both models, we computed indices that describe their respective performance in terms of predictive ability, discriminant ability, and overall fit. Finally, we re-estimated the models using combined data from 2005 and 2006 (“application sample”) and generated hospital risk standardized mortality rates and corresponding interval estimates.</p> <p>Model Development Dataset: The development sample consisted of 110,529 PCIs in 602 hospitals, with an overall unadjusted 30-day mortality rate of 1.4%.</p> <p>Model Performance: We computed 6 summary statistics for assessing model performance: over-fitting indices, percentage of variation explained by the risk factors, predictive ability, area under the receiver operating characteristic (ROC) curve, distribution of residuals, and model chi-square.</p> <p>The development model has excellent discrimination, calibration, and fit. The patient-level predicted mortality rate ranges from 0.1% in the lowest predicted decile to 7.0% in the highest predicted decile, a difference of 6.9%. The area under the ROC curve is 0.82. The discrimination and the explained variation of the model at the patient-level are consistent with those of published models of in-hospital PCI mortality (YNNH-CORE 2008).</p> <p>Model Validation: We compared the model performance in the development sample with its performance in a similarly derived validation sample from patients discharged in 2005 who had undergone PCI. This represented 88,630 cases discharged from the 457 hospitals in the 2005 validation dataset. This validation sample had a crude mortality rate (MR) of 1.4%. The standardized estimates and standard errors for the 2005 validation dataset are shown in Table 18 of the attached methodology report, and the performance metrics are shown in Table 19. The performance was not substantively different in this validation sample (ROC area = 0.82). As the results in Table 15 show, the 2005 and 2006 models appear well-calibrated.</p> <p>We examined the temporal variation of the standardized estimates and frequencies of the variables in the models (Tables 20 and 21). The frequencies and regression coefficients are consistent over the two years of data.</p>

	<p>Model Application: Table 22 in the methodology report shows the point estimates, standard errors, and associated T values for the HGLM for the full 2005-2006 application sample, calculated using the SAS GLIMMIX procedure. The estimated between-hospital variance in the adjusted log-odds of mortality is 0.1325 based on the full 2005-2006 dataset. This result implies that the odds of mortality for a high-mortality hospital (+1 SD) are 2.07 times that in a low-mortality hospital (-1 SD). If there were no differences between hospitals, the between-hospital variance would be 0 and the odds ratio would be 1.0.</p> <p>Testing Results: see above</p>
<p>27 (2d)</p>	<p>Measure Exclusions <i>Provide evidence to justify exclusion(s) and analysis of impact on measure results during testing.</i></p> <p>Summary of Evidence supporting exclusion(s):</p> <p>Citations for Evidence: We are using this field to list exclusions and to describe the rationale for each exclusion.</p> <p>Exclusions:</p> <p>1) Patients with >10 days between date of admission and date of PCI. Patients with with prolonged hospitalizations prior to PCI are excluded.</p> <p>Rationale for exclusion - The outcomes of patients with prolonged hospitalizations prior to PCI are less likely to be related to the PCI procedure.</p> <p>2) Transfer-in admissions (PCI to PCI). Among patients transferred from one acute care institution to another who had a PCI at both hospitals, the second admission with PCI is not eligible as an index admission. We used Medicare data to define transfers as two admissions that occur within 1 day of each other and identified patients in this cohort who had a PCI during both admissions.</p> <p>Rationale: We define an episode of care as starting on the first day of the first admission with PCI regardless of whether additional procedures are performed at the same hospital or at a different hospital after transfer.</p> <p>3) Admissions with missing death. Records with missing vital status in the Medicare enrollment file are excluded.</p> <p>Rationale: Records with no death information would prevent ascertainment of the outcome.</p> <p>4) Admissions which would lead to duplicate attribution of 30-day deaths.</p> <p>Rationale: The 30-day follow-up period for patients with more than one admission with PCI may overlap. In order to avoid attributing the same death to more than one admission with PCI (i.e. double counting a single patient death), later admissions with PCI were excluded.</p> <p>Data/sample: N/A</p> <p>Analytic Method: N/A</p> <p>Testing Results: N/A</p>
<p>28 (2e)</p>	<p>Risk Adjustment Testing <i>Summarize the testing used to determine the need (or no need) for risk adjustment and the statistical performance of the risk adjustment method.</i></p> <p>Data/sample: See Section # 26 above</p> <p>Analytic Method: Our approach to risk adjustment is tailored to and appropriate for a publicly reported outcome measure, as articulated in the AHA Scientific Statement, “Standards for Statistical Models Used for Public Reporting of Health Outcomes” (Krumholz et al., 2006).</p>

	<p>The development and validation datasets and samples are described above in # 26.</p> <p>Testing Results: In the development dataset, the ROC of 0.82 is higher than that of a model with just age and gender, 0.64, and the same as a model with all candidate variables, with ROC of 0.82.</p> <p>Adjusting for patient characteristics improved model performance.</p> <p>► If outcome or resource use measure not risk adjusted, provide rationale: N/A</p>
29 (2g)	<p>Testing comparability of results when more than 1 data method is specified (e.g., administrative claims or chart abstraction)</p> <p>Data/sample: N/A</p> <p>Analytic Method: N/A</p> <p>Results: N/A</p>
30 (2f)	<p>Provide Measure Results from Testing or Current Use (select one)</p> <p>Data/sample: N/A</p> <p>Methods to identify statistically significant and practically/meaningfully differences in performance: N/A</p> <p>Results: N/A</p>
31 (2h)	<p>Identification of Disparities</p> <p>► If measure is stratified by factors related to disparities (i.e. race/ethnicity, primary language, gender, SES, health literacy), provide stratified results: N/A</p> <p>► If disparities have been reported/identified, but measure is not specified to detect disparities, provide rationale: N/A</p>
USABILITY	
32 (3)	<p>Current Use <i>In development/testing</i> If in use, how widely used (select one) ► If "other," please describe: N/A</p> <p><input type="checkbox"/> Used in a public reporting initiative, name of initiative: Sample report attached <input type="checkbox"/> OR Web page URL:</p>
33 (3a)	<p>Testing of Interpretability (Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement)</p> <p>Data/sample: N/A</p> <p>Methods: Although there is no direct evidence that demonstrates the interpretability of the proposed measure, the methodology used to calculate this 30-day mortality measure parallels that used to calculate the 30-day mortality measures for AMI and HF, which were consumer-tested by CMS and are currently being publicly reported. In addition, similar measures are publicly reported in several states. Finally, in-hospital PCI mortality measures are currently used by the NCDR CathPCI Registry to benchmark hospital performance.</p> <p>Results: N/A</p>
34 (3b)	<p>Relation to other NQF-endorsed™ measures</p> <p>► Is this measure similar or related to measure(s) already endorsed by NQF (on the same topic or the same target population)? Measures can be found at www.qualityforum.org under Core Documents.</p>

3c)	<p><i>Check all that apply</i></p> <p><input type="checkbox"/> Have not looked at other NQF measures <input checked="" type="checkbox"/> Other measure(s) on same topic <input checked="" type="checkbox"/> Other measure(s) for same target population <input type="checkbox"/> No similar or related measures</p> <p>Name of similar or related NQF-endorsed™ measure(s): This measure complements three NQF-approved measures -- PCI in-hospital mortality measure, AMI 30-day mortality, HF 30-day mortality measure -- and a model submitted to NQF in this cycle for PCI patients with STEMI or shock. The pair of PCI mortality measures that are being submitted in this cycle, which were constructed in collaboration with the American College of Cardiology, are specifically designed for public reporting. The standardized period of follow-up, in contrast to the in-hospital model, is considered essential to a publicly reported measure so that all hospitals are judged equally, regardless of their length of stay or transfer policies. The in-hospital mortality measure was not intended for public reporting. Reporting this measure with AMI 30-day mortality and HF 30-day mortality is recommended so a hospital's performance across a range of cardiovascular conditions can be assessed.</p> <p>Are the measure specifications harmonized with existing NQF-endorsed™ measures? Yes, fully harmonized ▶ If not fully harmonized, provide rationale: The overall methodological approach for developing this measure parallels that used to develop the AMI, HF, and pneumonia 30-day mortality measures, which were previously approved by the National Quality Forum (NQF). The methodology is similar to that used by the National Cardiovascular Data Registry (NCDR) CathPCI in-hospital mortality model and utilizes similar variables for risk adjustment. However this model uses hierarchical modeling and stratifies PCI patients into two distinct cohorts that reflect their overall risk of procedural mortality.</p> <p>Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: The measure complements existing measures for 30-day mortality following admission for AMI or HF in that it will help provide a more complete picture of the outcomes achieved by hospitals across cardiovascular services. In addition, the measure adds value to the existing in-hospital NCDR PCI mortality model in that it is suitable for public reporting and will promote greater investment in quality improvement efforts.</p>
FEASIBILITY	
35 (4a)	<p><i>How are the required data elements generated? Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data elements are generated concurrent with and as a byproduct of care processes during care delivery (e.g., blood pressure or other assessment recorded by personnel conducting the assessment) <input type="checkbox"/> Data elements are generated from a patient survey (e.g., CAHPS) <input type="checkbox"/> Data elements are generated through coding performed by someone other than the person who obtained the original information (e.g., DRG or ICD-9 coding on claims) <input checked="" type="checkbox"/> Other, Please describe: <i>The outcome will be determined from an administrative database such as the Social Security Death Index.</i></p>
36 (4b)	<p><i>Electronic Sources All data elements</i></p> <p>▶ If all data elements are not in electronic sources, specify the near-term path to electronic collection by most providers:</p> <p>▶ Specify the data elements for the electronic health record:</p>
37 (4c)	<p><i>Do the specified exclusions require additional data sources beyond what is required for the other specifications? No</i></p> <p>▶ If yes, provide justification:</p>
38 (4d)	<p><i>Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure: Ensuring data quality is critical so that the RSMRs can provide fair and accurate estimates of outcomes across hospitals. However, all data sources are potentially prone to misclassifications. Accordingly, adequate mechanisms will need to be implemented to ensure data quality (such as monitoring data for variances in case mix [e.g., unexpectedly high proportion of salvage PCI or cardiogenic shock], chart audits, and</i></p>

	<p><i>possibly adjudicating cases that are vulnerable to systematic misclassification). The NCDR CathPCI registry has successfully implemented methods to ensure the quality of data used for the risk adjustment methodology, and a similar approach could be used by CMS when implementing this measure.</i></p> <p><i>Studies suggest that public reporting of the outcomes of cardiovascular procedures may have unintended consequences. Moscucci and colleagues compared the characteristics and outcomes of patients undergoing PCI in states with (New York) and without (Michigan) public reporting and found that patients undergoing PCI in New York were substantially lower risk than PCI patients in Michigan. Determining the underlying causes and appropriateness of these differences is impossible, but there is concern that physicians in states that publicly report PCI outcomes would either refer high risk cases to states without public reporting or avoid such cases altogether. Implementing a national measure of PCI outcomes would avoid the former problem in that public reporting would be consistent across states. Nevertheless, the proposed measure will require close attention to the possibility that high risk patients are not receiving PCI when clinically indicated. The proposed measure is, however, complementary to the previously approved measures for 30-day mortality of AMI and heart failure patients in that inappropriate avoidance of high risk PCI cases may have a detrimental effect on hospitals' performance on these other measures of cardiovascular outcomes.</i></p> <p><i>Describe how could these potential problems be audited: As discussed above, measure implementation will require close attention to data quality. Potential solutions include a) detailed chart audits, b) close attention to variances in case mix and c) review of some or all cases coded as cardiogenic shock or a salvage PCI</i></p> <p><i>Did you audit for these potential problems during testing? No If yes, provide results: N/A</i></p>
39 (4e)	<p><i>Testing feasibility Describe what have you learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: As noted previously, testing of this measure has not been performed. However, the NCDR CathPCI in-hospital mortality measure has already been approved by NQF. The implementation of this measure for benchmarking hospital performance demonstrates the feasibility of the proposed 30-day measures with regards to data collection, missing data, and data quality.</i></p>
CONTACT INFORMATION	
40	<p><i>Web Page URL for Measure Information Describe where users (implementers) should go for more details on specifications of measures, or assistance in implementing the measure. Web page URL: N/A</i></p>
41	<p>Measure Intellectual Property Agreement Owner Point of Contact First Name: MI: Last Name: Credentials (MD, MPH, etc.): Organization: Street Address: City: State: ZIP: Email: Telephone: ext:</p>
42	<p>Measure Submission Point of Contact If different than IP Owner Contact First Name: Lein MI: F Last Name: Han Credentials (MD, MPH, etc.): PhD Organization: Centers for Medicare & Medicaid Services (CMS) Street Address: 7500 Security Blvd City: Baltimore State: MD ZIP: 21244-9045 Email: Lein.Han@cms.hhs.gov Telephone: 410-786-0205 ext:</p>
43	<p>Measure Developer Point of Contact If different than IP Owner Contact First Name: Harlan MI: M Last Name: Krumholz Credentials (MD, MPH, etc.): MD Organization: Yale/YNHH Center for Outcomes Research and Evaluation (YNHH-CORE) Street Address: 1 Church Street, Suite 200 City: New Haven State: CT ZIP: 06510-3330 Email: harlan.krumholz@yale.edu Telephone: 203-764-9659 ext:</p>
44	<p>Measure Steward Point of Contact If different than IP Owner Contact <i>Identifies the organization that will take responsibility for updating the measure and assuring it is consistent with the scientific evidence and current coding schema; the steward of the measure may be</i></p>

	<p><i>different than the developer.</i> First Name: Lein MI:F Last Name:Han Credentials (MD, MPH, etc.): PhD Organization: Centers for Medicare & Medicaid Services (CMS) Street Address: 7500 Security Blvd City:Baltimore State:MD ZIP:21244-9045 Email: Lein.Han@cms.hhs.gov Telephone: 410-786-0205 ext</p>
ADDITIONAL INFORMATION	
45	<p>Workgroup/Expert Panel involved in measure development Workgroup/panel used ► If workgroup used, describe the members’ role in measure development: The measure developer, Yale/Yale-New Haven Hospital Center for Outcomes Research and Evaluation (YNHH-CORE) obtained expert and stakeholder input on the two measures through two mechanisms. First, the team has held regular conference calls with a Working Group of YNHH-CORE and American College of Cardiology (ACC)/National Cardiovascular Data Registry (NCDR) experts in cardiovascular registries and in the outcomes measure field. Second, YNHH-CORE sought and considered the input of an American College of Cardiology Foundation (ACCF) designated Task Force. ► Provide a list of workgroup/panel members’ names and organizations:</p> <p>Working Group</p> <p>Ralph Brindis, M.D., M.P.H., F.A.C.C. Regional Senior Advisor for Cardiovascular Disease, Northern California Kaiser Permanente; Clinical Professor of Medicine, UCSF, Oakland, CA; Chief Medical Officer and Chairman, Management Board, National Cardiovascular Data Registry</p> <p>Barbara Christensen, R.N., M.H.A. Senior Director, Registry Services, American College of Cardiology</p> <p>Jeptha Curtis, M.D. Assistant Professor of Medicine, Department of Internal Medicine (Cardiovascular Disease), Yale University</p> <p>Elizabeth Drye, M.D., S.M. Research Project Director, Yale/Yale-New Haven Hospital Center for Outcomes Research and Evaluation</p> <p>Susan Fitzgerald, R.N., M.B.A. Associate Director, Registry Development, American College of Cardiology</p> <p>Lori Geary, M.P.H. Research Project Coordinator, Yale/Yale-New Haven Hospital Center for Outcomes Research and Evaluation</p> <p>Amy Heller, Ph.D., M.P.H. Associate Director, Quality Products, American College of Cardiology</p> <p>Tony Hermann, R.N., M.B.A., C.P.H.Q. Associate Director, CathPCI Registry, American College of Cardiology</p> <p>Kathleen Hewitt, R.N., M.S.N., C.P.H.Q. Associate Vice President, American College of Cardiology</p> <p>Harlan Krumholz, M.D., M. Sc., F.A.C.C. Director, Yale Center for Outcomes Research and Evaluation; Representative, NCDR analytic center; Ex-officio to Task Force</p> <p>Kristi Mitchell, M.P.H. Senior Director, Research, Development and Quality Products, American College of Cardiology</p>

Eric Peterson, M.D., M.P.H., F.A.C.C.
Professor of Medicine, Duke University; Director, Cardiovascular Outcomes, Duke Clinical Research Institute, Chapel Hill, NC; Member, NCDR Science Oversight Committee/ Representative, NCDR Analytic Center

John Rumsfeld, M.D., Ph.D., F.A.C.C.
Associate Professor of Medicine, University of Colorado; Clinical Coordinator, VA Ischemic Heart Disease QUE, Denver, CO; Chief Science Officer, National Cardiovascular Data Registry

Lara Slattery, M.H.S.
Director, Quality Services Department - Registries, Products, and Publishing Division, American College of Cardiology

John Spertus, M.D., M.P.H., F.A.C.C.
Director of Cardiovascular Education and Outcomes Research, Mid America Heart Institute, Kansas City, MO; Member, NCDR Science Oversight Committee/Representative, NCDR analytic center; Chair, American College of Cardiology Foundation Task Force on Public Reporting of Hospital-Level Outcomes Measures

Yongfei Wang, M.S.
Senior Research Analyst, Yale/Yale-New Haven Hospital Center for Outcomes Research and Evaluation

William Weintraub, M.D., F.A.C.C.
Chair, CathPCI Registry Steering Committee; Section Chief, Cardiology, Christiana Care Health Services, Inc., Newark DE

Al Woodward, Ph.D., M.B.A.
Director, Research Services, American College of Cardiology

Task Force

Five Task Force members also serve as members of the Working Group, including:

Ralph G. Brindis, M.D., M.P.H., F.A.C.C.
Harlan Krumholz, M.D., M. Sc., F.A.C.C.
Eric Peterson, M.D., M.P.H., F.A.C.C.
John Rumsfeld, M.D., Ph.D., F.A.C.C.
John Spertus, M.D., M.P.H., F.A.C.C.

Other Task Force members are:

John Brush, M.D., F.A.C.C.
Cardiology Consultants LLC, Norfolk, VA; Chair, Quality Strategic Directions Committee

Vincent J. Bufalino, M.D., F.A.C.C.
Midwest Heart Specialists, Naperville, IL; Co-Chair, ACC Advocacy Committee

Gregory Dehmer, M.D., F.A.C.C.
Professor of Medicine, Texas A&M College of Medicine, Temple, TX;
Representative, The Society for Cardiovascular Angiography and Interventions

James Dove, M.D., F.A.C.C.
President, American College of Cardiology
President Emeritus, Prairie Cardiovascular Consultants, Ltd., Springfield, IL; President, ACC/ACCF Board of Trustees

Stephen C. Hammill, M.D., F.H.R.S.
Professor of Medicine, Mayo Clinic College of Medicine, Rochester, MN; Representative, Heart Rhythm

	<p>Society</p> <p>Frank E Harrell Jr., PhD Professor of Biostatistics; Department Chair, Vanderbilt University School of Medicine- Department of Biostatistics, Nashville, TN</p> <p>Barry K. Lewis, D.O., F.A.C.C. Consultants in Cardiology, P.C., Farmington Hills, MI; Member, Advocacy Committee</p> <p>William R. Lewis, M.D., F.A.C.C. Metro Health Medical Center, Cleveland, OH; ACC Ohio Chapter Governor/ACC Board of Governors</p> <p>Fred Masoudi, M.D., M.S.P.H., F.A.C.C. Denver Health Medical Center, Denver, CO; Chair, ACC/AHA Task Force on Performance Measures</p> <p>Andrea M. Russo, M.D. F.A.C.C. University of Pennsylvania Health System, Philadelphia, PA; Representative, Heart Rhythm Society</p> <p>Bonnie H. Weiner, M.D., F.S.C.A.I., F.A.C.C. Professor of Medicine; Interim Chair Cardiovascular Medicine, St. Vincent Hospital at Worcester Medical Center, Worcester, MA; Representative, The Society for Cardiovascular Angiography and Interventions</p> <p>Stuart Winston, D.O., F.A.C.C. Michigan Heart, P. C., Ann Arbor, MI; ACC Michigan Chapter Governor/ACC Board of Governors</p>
46	<p><i>Measure Developer/Steward Updates and Ongoing Maintenance</i></p> <p><i>Year the measure was first released: N/A</i></p> <p><i>Month and Year of most recent revision: N/A</i></p> <p><i>What is the frequency for review/update of this measure? N/A</i></p> <p><i>When is the next scheduled review/update for this measure? N/A</i></p>
47	<p>Copyright statement/disclaimers: N/A</p>
48	<p>Additional Information: Hospital 30-day Percutaneous Coronary Intervention Mortality Measures Methodology Report (attached)</p>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.0

August 2008

The measure information you submit will be shared with NQF's Steering Committees and Technical Advisory Panels to evaluate measures against the NQF criteria of importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Four conditions (as indicated below) must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, importance to measure and report, in order to be evaluated against the remaining criteria. References to the specific measure evaluation criteria are provided in parentheses following the item numbers. Please refer to the *Measure Evaluation Criteria* for more information at www.qualityforum.org under Core Documents. Additional guidance is being developed and when available will be posted on the NQF website.

Use the tab or arrow (↓→) keys to move the cursor to the next field (or back ←↑). There are three types of response fields:

- drop-down menus - select one response;
- check boxes - check as many as apply; and
- text fields - you can copy and paste text into these fields or enter text; these fields are not limited in size, but in most cases, we ask that you summarize the requested information.

Please note that URL hyperlinks do not work in the form; you will need to type them into your web browser.

Be sure to answer all questions. Fields that are left blank will be interpreted as no or none. Information must be provided in this form. Attachments are not allowed except when specifically requested or to provide additional detail or source documents for information that is summarized in this form. If you have important information that is not addressed by the questions, they can be entered into item #48 near the end of the form.

For questions about this form, please contact the NQF Project Director listed in the corresponding call for measures.

CONDITIONS FOR CONSIDERATION BY NQF	
	Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards.
A (A)	Public domain or Intellectual Property Agreement signed: Public domain - IP agreement not required (If no, do not submit) <i>Template for the Intellectual Property Agreement is available at www.qualityforum.org under Core Documents.</i>
B (B)	Measure steward/maintenance: Is there an identified responsible entity and process to maintain and update the measure on a schedule commensurate with clinical innovation, but at least every 3 years? Yes, information provided in contact section (If no, do not submit)
C (C)	Intended use: Does the intended use of the measure include BOTH public reporting AND quality improvement? Yes (If no, do not submit)
D (D)	Fully developed and tested: Is the measure fully developed AND tested? No, testing will be completed within 24 months (If not tested and no plans for testing within 24 months, do not submit)

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.0

August 2008

	(for NQF staff use) NQF Review #: HOE-010-08 NQF Project: Hospital Outcomes and Efficiency
MEASURE SPECIFICATIONS & DESCRIPTIVE INFORMATION	
1	Information current as of (date- MM/DD/YY): 11-21-08
2	Title of Measure: 30-day all-cause risk-standardized Percutaneous Coronary Intervention (PCI) mortality rate for patients with ST segment elevation myocardial infarction (STEMI) or cardiogenic shock
3	Brief description of measure ¹ : Hospital-specific 30-day all-cause risk-standardized mortality rate following Percutaneous Coronary Intervention (PCI) among patients aged 18 years or older with ST segment elevation myocardial infarction (STEMI) or cardiogenic shock at the time of procedure.
4 (2a)	<p>Numerator Statement: Note: This outcome measure does not have a traditional numerator and denominator like a core process measure (e.g., percentage of adult patients with diabetes aged 18-75 years receiving one or more hemoglobin A1c tests per year); thus, we use this field to define our statistically-adjusted rate outcome measure.</p> <p>We use hierarchical logistic regression modeling to calculate a hospital-specific 30-day risk-standardized mortality rate (RSMR). This rate is calculated as the ratio of “predicted” to “expected” deaths, multiplied by the national unadjusted mortality rate. For each hospital, the “numerator” of the ratio component of the RSMR is the predicted number of deaths within 30 days given the hospital’s performance with its observed case mix, and the “denominator” is the expected number of deaths given the hospital’s case mix. By convention, we use the term “predicted” here to describe the numerator result, which is calculated using the hospital-specific intercept term. We use “expected” for the denominator, which is calculated using the average intercept term.</p> <p>More specifically, the expected number of deaths for each hospital is estimated using its patient mix and the average hospital-specific intercept. The predicted number of deaths for each hospital is estimated given the same patient mix but the hospital-specific intercept. Operationally, the expected number of deaths for each hospital is obtained by regressing the risk factors (see Section #8) on the death using all hospitals in our sample, applying the subsequent estimated regression coefficients to the patient characteristics observed in the hospital, adding the average of the hospital-specific intercepts, transforming, and then summing over all patients in the hospital to get a value. This is a form of indirect standardization. The predicted hospital outcome is the number of deaths in the “specific” hospital estimated given its performance and case mix. Operationally, this is accomplished by estimating a hospital-specific intercept that represented baseline mortality risk within the hospital, applying the estimated regression coefficients to the patient characteristics in the hospital, transforming, and then summing over all patients in the hospital to get a value. To assess hospital performance in any given year, we re-estimate the model coefficients using that year’s data. (Please see the attached methodology report for details of the statistical methodology.)</p> <p>Time Window: This measure was developed with 24 months of data. The time period for public reporting has not been determined.</p> <p>Numerator Details (Definitions, codes with description):</p>
5 (2a)	<p>Denominator Statement: For “Denominator Statement,” please see #4 above. Instead, we are using this field to define our patient cohort.</p> <p>Outcome measure cohort definition: PCI procedures for patients at least 18 years of age, with STEMI or cardiogenic shock at the time of procedure.</p>

¹ Example of measure description: Percentage of adult patients with diabetes aged 18-75 years receiving one or more A1c test(s) per year.
NQF Measure Submission Form, V3.0

	<p>Time Window: This measure was developed with 24 months of data. The time period for public reporting has not been determined.</p> <p>Denominator Details (Definitions, codes with description): See above. We are using this field to specify the codes that define the PCI patient cohort</p> <p>In the CathPCI Registry, admissions with PCI are identified by field 614 (PCI=yes); STEMI or shock is defined as follows: (1) Symptoms present on admission = ACS:STEMI (field 550 = 6) with Time Period Symptom Onset to Admission within 24 hours (field 560 = 1,2,3) or Acute PCI = Yes (field 812 = 2,3,4); OR (2) Cardiogenic shock = Yes (field 520=1)</p>
<p>6 (2a, 2d)</p>	<p>Denominator Exclusions: Note: We are using this field to define exclusions to the patient cohort.</p> <p>(1) PCIs that follow a prior PCI in the same admission or occur during a transfer-in admission (PCI to PCI). We define an episode of care as starting on the day of the PCI during the first admission regardless of whether additional procedures are performed at the same hospital or at a different hospital after transfer. Thus, in the period of evaluation after the index procedure we do not begin a new period of evaluation after a second PCI during the same episode of care. If the patient is discharged to a non-acute care facility and has a second PCI within 30-days, that PCI is eligible as a new index PCI (except as noted in 3 below).</p> <p>(2) PCIs in patients with missing vital status (inability to link patient information to appropriate death index). In actual practice, with the identifiers that will be collected as part of the database we anticipate that missing data will be rare.</p> <p>(3) PCIs which would lead to duplicate attribution of 30-day deaths. The 30-day outcome period for patients with more than one PCI may overlap. In order to avoid attributing the same death to more than one PCI (i.e. double counting a single patient death), later PCI procedures within 30 days of the death are excluded.</p> <p>(4) PCIs for patients with more than 10 days between date of admission and date of PCI. Patients who have a PCI after many days of hospitalization are rare and represent a distinct population that likely has risk factors related to the hospitalization that are not well quantified in the registry. It seemed clinically sensible to exclude these patients.</p> <p>Denominator Exclusion Details (Definitions, codes with description): See above. We are deriving the corresponding codes based on the data for exclusion.</p>
<p>7 (2a, 2h)</p>	<p>Stratification Do the measure specifications require the results to be stratified? No ► If “other” describe:</p> <p>Identification of stratification variable(s):</p> <p>Stratification Details (Definitions, codes with description): This measure was designed to be reported along with 30-day all-cause risk-standardized percutaneous coronary intervention (PCI) mortality rate for patients without ST segment elevation myocardial infarction (STEMI) and without cardiogenic shock</p>
<p>8 (2a, 2e)</p>	<p>Risk Adjustment Does the measure require risk adjustment to account for differences in patient severity before the onset of care? Yes ► If yes, Statistical Risk Model, see Variables ► Is there a separate proprietary owner of the risk model? No</p> <p>Identify Risk Adjustment Variables: Age (10 year increments) Body Mass Index (5 kg/m² increments) Cerebrovascular disease Chronic Lung disease Glomerular Filtration Rate (GFR) (derived)</p>

	<p>0=Not measured 1="GFR<30" 2="30≤GFR<60" 3="60≤GFR<90" 4="GFR≥90"</p> <p>Previous PCI Heart Failure - Current Status Cardiogenic shock on admission Symptom Onset No MI on admission MI within 24 hours of admission MI > 24 hours after admission</p> <p>Ejection Fraction Percent (EF) 1=Not measured 2="EF<30" 3="30≤ EF<45" 4="EF≥45"</p> <p>PCI status 1=Elective 2=Urgent 3=Emergency 4=Salvage</p> <p>Highest Risk Lesion - coronary artery segment category 1=proximal Right Coronary Artery (RCA)/mid Left Anterior Descending (LAD) artery/proximal Circumflex Artery (Cx) 2=proximal LAD 3=Left Main</p> <p>Highest Risk Lesion: Society for Cardiovascular Angiography and Interventions (SCAI) class 1 class 2 or 3 class 4</p> <p>For more details, please see the attached methodology report.</p> <p>Detailed risk model: attached <input checked="" type="checkbox"/> OR Web page URL:</p>
<p>9</p> <p>(2a)</p>	<p>Type of Score: Rate/proportion Calculation Algorithm: attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p><i>(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)</i></p> <p>Other ► If "Other", please describe: For each hospital, we calculate a 30-day all-cause risk-standardized mortality rate (RSMR) and an interval estimate for the RSMR, which expresses the level of uncertainty around the point estimate. The RSMR with its interval estimate can be used to classify hospital performance (e.g., higher than expected, as expected, or lower than expected).</p>
<p>10</p> <p>(2a, 4a, 4b)</p>	<p>Identify the required data elements(e.g., primary diagnosis, lab values, vital signs): Data dictionary/code table attached <input type="checkbox"/> OR Web page URL: http://www.ncdr.com/WebNCDR/ELEMENTS.ASPX</p> <p>Data Quality (2a) <i>Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data are captured from an authoritative/accurate source (e.g., lab values from laboratory personnel) <input checked="" type="checkbox"/> Data are coded using recognized data standards <input type="checkbox"/> Method of capturing data electronically fits the workflow of the authoritative source <input type="checkbox"/> Data are available in EHRs <input checked="" type="checkbox"/> Data are auditable</p>
<p>11</p> <p>(2a,</p>	<p>Data Source and Data Collection Methods <i>Identifies the data source(s) necessary to implement the measure specifications. Check all that apply</i></p> <p><input type="checkbox"/> Electronic Health/Medical Record <input type="checkbox"/> Paper Medical Record</p>

4b)	<input type="checkbox"/> Electronic Clinical Database, Name: <input checked="" type="checkbox"/> Electronic Clinical Registry, Name: National Cardiovascular Data Registry, CathPCI Registry; required data could alternatively be collected through other non-registry mechanisms <input type="checkbox"/> Electronic Claims <input type="checkbox"/> Electronic Pharmacy data <input type="checkbox"/> Electronic Lab data <input type="checkbox"/> Electronic source - other, Describe:	<input type="checkbox"/> Standardized clinical instrument, Name: <input type="checkbox"/> Standardized patient survey, Name: <input type="checkbox"/> Standardized clinician survey, Name: <input checked="" type="checkbox"/> Other, Describe: Death Index Instrument/survey attached <input checked="" type="checkbox"/> OR Web page URL:
12 (2a)	Sampling <i>If measure is based on a sample, provide instructions and guidance on sample size. Minimum sample size: Data from all hospitals and all PCI would be included in the process of reestimating model variables. For public reporting, minimum sample size has not been determined.</i> Instructions: N/A	
13 (2a)	Type of Measure: Outcome ▶ If “Other”, please describe: ▶ If part of a composite or paired with another measure, please identify composite or paired measure This measure is being submitted along with: paired with 30-day all-cause risk-standardized percutaneous coronary intervention (PCI) mortality rate for patients without ST segment elevation myocardial infarction (STEMI) and without cardiogenic shock	
14 (2a)	Unit of Measurement/Analysis <i>(Who or what is being measured) Check all that apply.</i> <input type="checkbox"/> Can be measured at all levels <input type="checkbox"/> Individual clinician (e.g., physician, nurse) <input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice) <input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home) <input type="checkbox"/> Integrated delivery system <input type="checkbox"/> Health plan <input type="checkbox"/> Community/Population <input type="checkbox"/> Other <i>(Please describe):</i>	
15 (2a)	Applicable Care Settings <i>Check all that apply</i> <input type="checkbox"/> Can be used in all healthcare settings <input type="checkbox"/> Ambulatory Care (office/clinic) <input type="checkbox"/> Behavioral Healthcare <input type="checkbox"/> Community Healthcare <input type="checkbox"/> Dialysis Facility <input type="checkbox"/> Emergency Department <input type="checkbox"/> EMS emergency medical services <input type="checkbox"/> Health Plan <input type="checkbox"/> Home Health <input type="checkbox"/> Hospice <input checked="" type="checkbox"/> Hospital <input type="checkbox"/> Long term acute care hospital <input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF) <input type="checkbox"/> Prescription Drug Plan <input type="checkbox"/> Rehabilitation Facility <input type="checkbox"/> Substance Use Treatment Program/Center <input type="checkbox"/> Other <i>(Please describe):</i>	
IMPORTANCE TO MEASURE AND REPORT		
Note: This is a threshold criterion. If a measure is not judged to be sufficiently important to measure and report, it will not be evaluated against the remaining criteria.		
16 (1a)	Addresses a Specific National Priority Partners Goal <i>Enter the numbers of the specific goals related to this measure (see list of goals on last page):</i> 3.3, 3.4	
17 (1a)	If not related to NPP goal, identify high impact aspect of healthcare (select one) Summary of Evidence: Citations² for Evidence:	
18 (1b)	Opportunity for Improvement <i>Provide evidence that demonstrates considerable variation, or overall poor performance, across providers.</i> Summary of Evidence: PCI is one of the most commonly performed cardiac procedures in the United	

² Citations can include, but are not limited to journal articles, reports, web pages (URLs).
 NQF Measure Submission Form, V3.0

	<p>States. In 2005, an estimated 1,265,000 PCI procedures were performed in the United States (Rosamond Flegal et al. 2008). From 1987-2003, the number of procedures increased 326% (Thom, Haase et al. 2006). Inpatient mortality is the indicator that has been most widely used to evaluate the quality of cardiac procedures and is arguably the most important adverse outcome measure. The ACC summarized the experience of the NCDR CathPCI Registry from 1998-2000 and found that inhospital mortality occurred in 1,422 of 100,253 PCI procedures (1.4%) (Shaw, Anderson et al. 2002). Mortality was higher in patients with acute myocardial infarction (4.9%) or cardiogenic shock (27.2%). In the present era, mortality rates for PCI in large series from experienced operators varied across hospitals (Carrozza, Cutlip et al. 2008). Prior studies have demonstrated significant variability in in-hospital PCI mortality across age groups, gender, geographic regions, socioeconomic status, and by hospital volume (Mukherjee, Wainess et al. 2005). Although 12 states already report PCI outcomes, to date there has not been a unified national effort to publicly report PCI mortality.</p> <p>Citations for Evidence: Carrozza J, Cutlip D, Levin T. (2008). Periprocedural complications of percutaneous coronary intervention. UpToDate. B. Rose. Waltham, MA.</p> <p>Mukherjee D, Wainess RM, et al. (2005). "Variation in outcomes after percutaneous coronary intervention in the United States and predictors of periprocedural mortality." <i>Cardiology</i> 103(3): 143-7.</p> <p>Rosamond W, Flegal K, Furie K, Go A, Greenlund K, Haase N, Hailpern SM, Ho M, Howard V, Kissela B, Kittner S, Lloyd-Jones D, McDermott M, Meigs J, Moy C, Nichol G, O'Donnell C, Roger V, Sorlie P, Steinberger J, Thom T, Wilson M, Hong Y. Heart Disease and Stroke Statistics_2008 Update: A Report From the American Heart Association Statistics Committee and Stroke Statistics Subcommittee and for the American Heart Association Statistics Committee and Stroke Statistics Subcommittee <i>Circulation</i> 2008;117:e25-e146; originally published online Dec 17, 2007; DOI: 10.1161/CIRCULATIONAHA.107.187998.</p> <p>Shaw RE, Anderson HV, et al. (2002). "Development of a risk adjustment mortality model using the American College of Cardiology-National Cardiovascular Data Registry (ACC-NCDR) experience: 1998-2000." <i>J Am Coll Cardiol</i> 39(7): 1104-12.</p> <p>Thom T, Haase N, et al. (2006). "Heart disease and stroke statistics--2006 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee." <i>Circulation</i> 113(6): e85-151.</p>
<p>19 (1b)</p>	<p>Disparities <i>Provide evidence that demonstrates disparity in care/outcomes related to the measure focus among populations.</i></p> <p>Summary of Evidence: We have not examined health disparities associated with this measure. This measure could be used to assess differences in performance among hospitals that care for different types of populations (e.g., those that serve primarily minority populations versus others).</p> <p>Citations for evidence: N/A</p>
<p>20 (1c)</p>	<p>If measuring an Outcome Describe relevance to the national health goal/priority, condition, population, and/or care being addressed: This measure will describe hospital-level mortality rates following PCI with the overriding goal to reduce preventable and premature mortality rates to best-in-class (NPP 3.3) and 30-day mortality rates following hospitalization for relevant conditions to best-in-class (NPP 3.4).</p> <p>If not measuring an outcome, provide evidence supporting this measure topic and grade the strength of the evidence <i>Summarize the evidence (including citations to source) supporting the focus of the measure as follows:</i></p> <ul style="list-style-type: none"> • Intermediate outcome - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit. • Process - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s). • Structure - evidence that the measured structure supports the consistent delivery of effective

- processes or access that lead to improved health/avoidance of harm or cost/benefit.
- Patient experience - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
 - Access - evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
 - Efficiency - demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

Type of Evidence *Check all that apply*

- | | |
|---|--|
| <input type="checkbox"/> Evidence-based guideline | <input type="checkbox"/> Quantitative research studies |
| <input type="checkbox"/> Meta-analysis | <input type="checkbox"/> Qualitative research studies |
| <input type="checkbox"/> Systematic synthesis of research | <input type="checkbox"/> Other (<i>Please describe</i>): |

Overall Grade for Strength of the Evidence³ (*Use the USPSTF system, or if different, also describe how it relates to the USPSTF system*): **N/A**

Summary of Evidence (*provide guideline information below*):

Evidence that the outcome measure has been influenced by one or more clinical interventions:

Numerous studies have demonstrated the efficacy of interventions designed to improve patient outcomes following PCI. These include pharmacologic interventions such as the use of glycoprotein 2b/3a inhibitors, direct thrombin inhibitors, and pre-procedural clopidogrel, as well as advances in device technology such as use of stents (and more recently drug eluting stents), thrombectomy for acute lesions with high thrombus burden, and distal embolic protection for PCI of degenerated saphenous vein grafts. Of note, the majority of these interventions have been shown to reduce endpoints other than mortality, most commonly rates of periprocedural MI, major bleeding, and target vessel revascularization for in-stent restenosis. Although few individual interventions have been shown to reduce mortality, they may collectively exert a favorable impact on hospital PCI mortality rates when implemented in a coordinated fashion.

There is a growing body of evidence that quality improvement efforts can improve outcomes of PCI patients, including survival. Rihal and colleagues examined patient outcomes before and after initiation of a program of continuous quality improvement (CQI) and found a significantly lower in-hospital mortality following PCI despite significant increases in the risk profile of PCI patients. Similar improvements were identified in studies of CQI by Brush et al and Moscucci et al, and improvements in survival were associated with greater adherence to evidence based practices including preprocedural clopidogrel, use of glycoprotein 2b/3a inhibitors, and volume of iodinated contrast. The observational nature of these studies precludes drawing definitive conclusions, but they strongly suggest a mechanism by which public reporting of hospital PCI outcomes could promote improvements in the care of PCI patients.

Citations for Evidence:

Brush JE, Balakrishnan SA, Brough J, Hartman C, Hines G, Liverman DP, Parker JP, Rich J, Tindall N. (2006). "Implementation of a continuous quality improvement program for percutaneous coronary intervention and cardiac surgery at a large community hospital." *Am Heart J* 152 (2):379-85 16875926 (P,S,E,B).

Moscucci M, Kline Rogers E, Montoye C, Smith DE, Share D, O'Donnell M, Maxwell-Eward A, Meengs WL, De Franco AC, Patel K, McNamara R, McGinnity JG, Jani SM, Khanal S, Eagle KA. (2006). "Association of a Continuous Quality Improvement Initiative With Practice and Outcome Variations of Contemporary Percutaneous Coronary Interventions." *Circulation*. 113:814-822.

Rihal C, Kamath C, Holmes D, et al. (2006). "Economic and clinical outcomes of a physician-led continuous quality improvement intervention in the delivery of percutaneous coronary intervention." *Am J Manag Care* 12:445-452.

³The strength of the body of evidence for the specific measure focus should be systematically assessed and rated, e.g., USPSTF grading system www.ahrq.gov/clinic/uspstmeth.htm: A - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. B - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. C - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if

NQF Measure Submission Form, V3.0

<p>21 (1c)</p>	<p>Clinical Practice Guideline <i>Cite the guideline reference; quote the specific guideline recommendation related to the measure and the guideline author's assessment of the strength of the evidence; and summarize the rationale for using this guideline over others.</i></p> <p>Guideline Citation: N/A</p> <p>Specific guideline recommendation: N/A</p> <p>Guideline author's rating of strength of evidence (<i>If different from USPSTF, also describe it and how it relates to USPSTF</i>): N/A</p> <p>Rationale for using this guideline over others: N/A</p>
<p>22 (1c)</p>	<p>Controversy/Contradictory Evidence <i>Summarize any areas of controversy, contradictory evidence, or contradictory guidelines and provide citations.</i></p> <p>Summary: There are a few points of potential controversy that deserve comment:</p> <p>1) This model is designed for use in national public reporting and is aligned with the American Heart Association (AHA) published standards for publicly reported outcomes measures (Krumholz et al. 2006). The model, however, was developed from a subset of the entire population of PCI patients, namely fee-for-service Medicare patients undergoing PIC at facilities that participate in the NCDR CathPCI Registry. Furthermore, patients' vital status at 30 days was determined by linking to administrative data using a probabilistic match because our derivation and validation samples lacked unique identifiers to merge with a national death index.</p> <p>For public reporting, the parameters would be re-estimated using the national data. In addition, direct identifiers would be used to link clinical data and determine vital status. Further, adequate mechanisms would need to be implemented to ensure data quality (such as monitoring data for variances in case mix [e.g., unexpectedly high proportion of salvage PCI or cardiogenic shock], chart audits, and possibly adjudicating cases that are vulnerable to systematic misclassification). There is no reason to believe that these changes will significantly change the performance of the model.</p> <p>2) We chose 30-day mortality as our period of assessment. In contrast, prior efforts to create models for risk adjusting PCI outcomes have used hospitals' self-reported in-hospital mortality. Advantages of a 30-day mortality outcome include providing a standardized period of assessment, potentially more accurate assessments of vital status, and providing a more complete picture of outcomes following PCI. The main disadvantage of this approach is that it requires linking clinical data to a separate data source to determine 30-day vital status. In order to inform this decision, we determined whether in-hospital mortality rates were comparable to 30-day mortality rates. We found that the median absolute difference in unadjusted mortality rates was 0.5%, but that 8% of hospitals had a greater than 2% absolute difference. Furthermore, when we compared hospitals' decile ranking using these two endpoints, a quarter of hospitals changed more than one decile. Based on this evidence, we determined that in-hospital mortality may not be an adequate surrogate for 30-day mortality following PCI.</p> <p>3) We propose models that stratify the population of patients undergoing PCI into two distinct cohorts: patients with STEMI or cardiogenic shock, and patients without STEMI and without shock. This approach reflects the fact that among patients undergoing PCI, the risk of mortality differs considerably depending on the clinical context in which it is performed. The mortality of PCI patients with an evolving STEMI is substantially higher than that of outpatients undergoing elective procedures. Furthermore, many hospitals (e.g., primary PCI centers) can only perform PCI on patients undergoing non-elective procedures. Stratifying the population into these cohorts was felt to provide more fair and accurate comparisons of the outcomes of patients treated at different types of hospitals. This strategy has previously been</p>

other considerations support the offering or providing the service in an individual patient. D - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. I - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

	<p>implemented by the Massachusetts program for publicly reporting of mortality following PCI (www.massdac.org/pic/index.htm). The state of New York reports outcomes for both the combined cohort as well as a stratified cohort (www.health.state.ny.us/statistics/diseases/cardiovascular).</p> <p>4) We did not consider as candidate variables those that we would not want to adjust for in a quality measure, such as potential complications, certain patient demographics (e.g., race, socioeconomic status), and patients' admission path (e.g., admitted from, or discharged to, a skilled nursing facility [SNF]). These characteristics may be associated with mortality and thus could increase the model performance to predict patient mortality. However, these variables may be related to quality or supply factors that should not be included in an adjustment that seeks to control for patient clinical characteristics while illuminating important quality differences.</p> <p>5) Studies suggest that public reporting of the outcomes of cardiovascular procedures may have unintended consequences. Moscucci and colleagues compared the characteristics and outcomes of patients undergoing PCI in states with (New York) and without (Michigan) public reporting and found that patients undergoing PCI in New York were substantially lower risk than PCI patients in Michigan. Determining the underlying causes and appropriateness of these differences is impossible, but there is concern that physicians in states that publicly report PCI outcomes would either refer high risk cases to states without public reporting or avoid such cases altogether. Implementing a national measure of PCI outcomes would avoid the former problem in that public reporting would be consistent across states. Nevertheless, the proposed measure will require close attention to the possibility that high risk patients are not receiving PCI when clinically indicated. The proposed measure is, however, complementary to the previously approved measures for 30-day mortality of AMI and heart failure patients in that inappropriate avoidance of high risk PCI cases may have a detrimental effect on hospitals' performance on these other measures of cardiovascular outcomes.</p> <p>Citations: Krumholz HM, Brindis RG, et al. (2006). "Standards for statistical models used for public reporting of health outcomes: an American Heart Association Scientific Statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council. Endorsed by the American College of Cardiology Foundation." <i>Circulation</i> 113(3): 456-62.</p>
23 (1)	<p>Briefly describe how this measure (as specified) will facilitate significant gains in healthcare quality related to the specific priority goals and quality problems identified above: Public reporting will drive internal hospital quality improvement efforts to achieve mortality rates consistent with or better than the best performing hospitals in the country.</p>
SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
<p>Note: Testing and results should be summarized in this form. However, additional detail and reports may be submitted as supplemental information or provided as a web page URL. If a measure has not been tested, it is only potentially eligible for time-limited endorsement.</p>	
24	<p>Supplemental Testing Information: attached <input checked="" type="checkbox"/> OR Web page URL:</p>
25 (2b)	<p>Reliability Testing</p> <p>Data/sample: <i>N/A</i></p> <p>Analytic Method: <i>N/A</i></p> <p>Testing Results: Reliability testing using data from all PCIs with STEMI or Shock will be performed prior to measure implementation.</p>
26 (2c)	<p>Validity Testing</p> <p>Data/sample: We are using this field to describe model testing and validation.</p> <p>We developed this model in a limited cohort (see Section 22 above).</p>

We developed a model in a cohort of NCDR patients ≥ 65 years old who had matching information in Medicare administrative data that allowed us to link to the outcome (mortality within 30 days). In the Medicare dataset, admissions with PCI are identified by International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) procedure codes, shown here:

- 00.66 Percutaneous transluminal coronary angioplasty or coronary atherectomy
- 36.01 Single vessel PTCA or coronary atherectomy
- 36.02 Percutaneous transluminal coronary angioplasty or coronary atherectomy with mention of thrombolytic agent
- 36.05 Multiple vessel PTCA or coronary atherectomy
- 36.06 Insertion of non-drug-eluting coronary artery stent(s)
- 36.07 Insertion of drug-eluting coronary artery stent(s)

Analytic Method:

Because patient identifiers are not currently available, we used a probabilistic match to link the two datasets using hospital Medicare Provider Number (MPN), patient age, gender, date of admission, and date of discharge. We matched 65% of the PCI cohort after excluding non-unique records. Matched and unmatched patients had similar clinical characteristics.

Overview: The evidence supporting the measure can be found in the Methodology report accompanying this submission. In brief, a risk adjustment model was derived using all matched admissions in 2006 (“development sample”). The performance of the models was validated using a similar cohort of patients who underwent PCI in 2005 (“validation sample”). For both models, we computed indices that describe their respective performance in terms of predictive ability, discriminant ability, and overall fit. Finally, we re-estimated the models using combined data from 2005 and 2006 (“application sample”) and generated hospital risk-standardized mortality rates and corresponding interval estimates.

Model Development Dataset: The development sample consisted of 15,123 PCIs in 602 hospitals, with an overall unadjusted 30-day mortality rate of 9.2%.

Model Performance: We computed 6 summary statistics for assessing model performance: over-fitting indices, percentage of variation explained by the risk factors, predictive ability, area under the receiver operating characteristic (ROC) curve, distribution of residuals, and model chi-square.

The development model has excellent discrimination, calibration, and fit. The patient-level mortality rate ranges from 1.4% in the lowest predicted decile to 40.3% in the highest predicted decile, with a difference of 38.9%. The area under the ROC curve is 0.83. The discrimination and the explained variation of the model at the patient-level are consistent with those of published PCI in-hospital mortality models (YNHH-CORE 2008).

Model Validation: We compared the model performance in the development sample with its performance in a similarly derived sample from patients discharged in 2005 who had undergone PCI. There were 12,052 cases discharged from the 458 hospitals in the 2005 validation dataset. This validation sample had a crude mortality rate of 9.0%. The standardized estimates and standard errors for the 2005 validation dataset are shown in Table 12 of the attached methodology report, and the performance metrics are shown in Table 13. The performance was not substantively different in this validation sample (ROC = 0.84), as compared to the development sample (ROC = 0.83). As the results in Table 9 show, the 2005 and 2006 models are similarly calibrated.

We examined the temporal variation of the standardized estimates and frequencies of the variables in the models (Tables 14 and 15). The frequencies and regression coefficients are consistent over the two years of data.

Model Application: Table 16 in the methodology report shows the point estimates, standard errors, and associated T values for the HGLM for the 2005-2006 combined dataset, calculated using the SAS GLIMMIX procedure. The estimated between-hospital variance in the adjusted log-odds of mortality is 0.1024, based on the 2005-2006 combined dataset. This result implies that the odds of mortality for a high-

	<p>mortality hospital (+1 SD) are 1.90 times that in a low-mortality hospital (-1 SD). If there were no differences between hospitals, the between-hospital variance would be 0 and the odds ratio would be 1.0.</p> <p>Testing Results: <i>see above</i></p>
<p>27 (2d)</p>	<p>Measure Exclusions <i>Provide evidence to justify exclusion(s) and analysis of impact on measure results during testing.</i></p> <p>Summary of Evidence supporting exclusion(s):</p> <p>Citations for Evidence: We are using this field to list exclusions and to describe the rationale for each exclusion.</p> <p>Exclusions:</p> <p>1) Patients with >10 days between date of admission and date of PCI. Patients with with prolonged hospitalizations prior to PCI are excluded.</p> <p>Rationale for exclusion - The outcomes of patients with prolonged hospitalizations prior to PCI are less likely to be related to the PCI procedure.</p> <p>2) Transfer-in admissions (PCI to PCI). Among patients transferred from one acute care institution to another who had a PCI at both hospitals, the second admission with PCI is not eligible as an index admission. We used Medicare data to define transfers as two admissions that occur within 1 day of each other and identified patients in this cohort who had a PCI during both admissions.</p> <p>Rationale: We define an episode of care as starting on the first day of the first admission with PCI regardless of whether additional procedures are performed at the same hospital or at a different hospital after transfer.</p> <p>3) Admissions with missing death. Records with missing vital status in the Medicare enrollment file are excluded.</p> <p>Rationale: Records with no death information would prevent ascertainment of the outcome.</p> <p>4) Admissions which would lead to duplicate attribution of 30-day deaths.</p> <p>Rationale: The 30-day follow-up period for patients with more than one admission with PCI may overlap. In order to avoid attributing the same death to more than one admission with PCI (i.e. double counting a single patient death), later admissions with PCI were excluded.</p> <p>Data/sample: N/A</p> <p>Analytic Method: N/A</p> <p>Testing Results: N/A</p>
<p>28 (2e)</p>	<p>Risk Adjustment Testing <i>Summarize the testing used to determine the need (or no need) for risk adjustment and the statistical performance of the risk adjustment method.</i></p> <p>Data/sample: <i>See Section # 26 above</i></p> <p>Analytic Method: Our approach to risk adjustment is tailored to and appropriate for a publicly reported outcome measure, as articulated in the AHA Scientific Statement, “Standards for Statistical Models Used for Public Reporting of Health Outcomes” (Krumholz et al., 2006).</p> <p>The development and validation datasets and samples are described above in Section # 26.</p> <p>Testing Results: In the development dataset, the ROC of 0.83 is higher than that of a model with just age and gender, 0.62, and the same as a model with all candidate variables, with ROC of 0.83.</p>

	<p>Adjusting for patient characteristics improved model performance.</p> <p>► If outcome or resource use measure not risk adjusted, provide rationale: <i>N/A</i></p>
29 (2g)	<p>Testing comparability of results when more than 1 data method is specified (<i>e.g., administrative claims or chart abstraction</i>)</p> <p>Data/sample: <i>N/A</i></p> <p>Analytic Method: <i>N/A</i></p> <p>Results: <i>N/A</i></p>
30 (2f)	<p>Provide Measure Results from Testing or Current Use (<i>select one</i>)</p> <p>Data/sample: <i>N/A</i></p> <p>Methods to identify statistically significant and practically/meaningfully differences in performance: <i>N/A</i></p> <p>Results: <i>N/A</i></p>
31 (2h)	<p>Identification of Disparities</p> <p>► If measure is stratified by factors related to disparities (i.e. race/ethnicity, primary language, gender, SES, health literacy), provide stratified results: <i>N/A</i></p> <p>► If disparities have been reported/identified, but measure is not specified to detect disparities, provide rationale: <i>N/A</i></p>
USABILITY	
32 (3)	<p>Current Use <i>In development/testing</i> If in use, how widely used (<i>select one</i>) ► If “other,” please describe: <i>N/A</i></p> <p><input type="checkbox"/> Used in a public reporting initiative, name of initiative: Sample report attached <input type="checkbox"/> OR Web page URL:</p>
33 (3a)	<p>Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)</p> <p>Data/sample: <i>N/A</i></p> <p>Methods: <i>Although there is no direct evidence that demonstrates the interpretability of the proposed measure, the methodology used to calculate this 30-day mortality measure parallels that used to calculate the 30-day mortality measures for AMI and HF, which were consumer-tested by CMS and are currently being publicly reported. In addition, similar measures are publicly reported in several states. Finally, in-hospital PCI mortality measures are currently used by the NCDR CathPCI Registry to benchmark hospital performance.</i></p> <p>Results: <i>N/A</i></p>
34 (3b, 3c)	<p>Relation to other NQF-endorsed™ measures</p> <p>► Is this measure similar or related to measure(s) already endorsed by NQF (on the same topic or the same target population)? <i>Measures can be found at www.qualityforum.org under Core Documents. Check all that apply</i></p> <p><input type="checkbox"/> Have not looked at other NQF measures <input checked="" type="checkbox"/> Other measure(s) on same topic <input checked="" type="checkbox"/> Other measure(s) for same target population <input type="checkbox"/> No similar or related measures</p> <p>Name of similar or related NQF-endorsed™ measure(s): <i>This measure complements three NQF approved measures -- PCI in-hospital mortality measure, AMI 30-day mortality, HF 30-day mortality measure -- and a model submitted to NQF in this cycle for PCI patients</i></p>

	<p>without STEMI and without shock. The pair of PCI mortality measures that are being submitted in this cycle, which were constructed in collaboration with the American College of Cardiology, are specifically designed for public reporting. The standardized period of follow-up, in contrast to the in-hospital model, is considered essential to a publicly reported measure so that all hospitals are judged equally, regardless of their length of stay or transfer policies. The in-hospital mortality measure was not intended for public reporting. Reporting this measure with AMI 30-day mortality and HF 30-day mortality is recommended so a hospital's performance across a range of cardiovascular conditions can be assessed.</p> <p>Are the measure specifications harmonized with existing NQF-endorsed™ measures? Yes, fully harmonized ► If not fully harmonized, provide rationale: The overall methodological approach for developing this measure parallels that used to develop the AMI, HF, and pneumonia 30-day mortality measures, which were previously approved by the National Quality Forum (NQF). The methodology is similar to that used by the National Cardiovascular Data Registry (NCDR) CathPCI in-hospital mortality model and uses similar variables. However this model uses hierarchical modeling and stratifies PCI patients into two distinct cohorts that reflect their overall risk of procedural mortality.</p> <p>Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: The measure complements existing measures for 30-day mortality following admission for AMI or HF in that it will help provide a more complete picture of the outcomes achieved by hospitals across cardiovascular services. In addition, the measure adds value to the existing in-hospital NCDR PCI mortality model in that it is suitable for public reporting and will promote greater investment in quality improvement efforts.</p>
FEASIBILITY	
<p>35 (4a)</p>	<p>How are the required data elements generated? <i>Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data elements are generated concurrent with and as a byproduct of care processes during care delivery (e.g., blood pressure or other assessment recorded by personnel conducting the assessment)</p> <p><input type="checkbox"/> Data elements are generated from a patient survey (e.g., CAHPS)</p> <p><input type="checkbox"/> Data elements are generated through coding performed by someone other than the person who obtained the original information (e.g., DRG or ICD-9 coding on claims)</p> <p><input checked="" type="checkbox"/> Other, Please describe: <i>The outcome will be determined from an administrative database such as the Social Security Death Index.</i></p>
<p>36 (4b)</p>	<p>Electronic Sources <i>All data elements</i></p> <p>► If all data elements are not in electronic sources, specify the near-term path to electronic collection by most providers:</p> <p>► Specify the data elements for the electronic health record:</p>
<p>37 (4c)</p>	<p>Do the specified exclusions require additional data sources beyond what is required for the other specifications? <i>No</i></p> <p>► If yes, provide justification:</p>
<p>38 (4d)</p>	<p>Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure: <i>Ensuring data quality is critical so that the RSMRs can provide fair and accurate estimates of outcomes across hospitals. However, all data sources are potentially prone to misclassifications. Accordingly, adequate mechanisms will need to be implemented to ensure data quality (such as monitoring data for variances in case mix [e.g., unexpectedly high proportion of salvage PCI or cardiogenic shock], chart audits, and possibly adjudicating cases that are vulnerable to systematic misclassification). The NCDR CathPCI registry has successfully implemented methods to ensure the quality of data used for the risk adjustment methodology, and a similar approach could be used by CMS when implementing this measure.</i></p> <p><i>Studies suggest that public reporting of the outcomes of cardiovascular procedures may have unintended consequences. Moscucci and colleagues compared the characteristics and outcomes of patients undergoing PCI in states with (New York) and without (Michigan) public reporting and found that patients undergoing PCI in New York were substantially lower risk than PCI patients in Michigan. Determining the underlying</i></p>

	<p>causes and appropriateness of these differences is impossible, but there is concern that physicians in states that publicly report PCI outcomes would either refer high risk cases to states without public reporting or avoid such cases altogether. Implementing a national measure of PCI outcomes would avoid the former problem in that public reporting would be consistent across states. Nevertheless, the proposed measure will require close attention to the possibility that high risk patients are not receiving PCI when clinically indicated. The proposed measure is, however, complementary to the previously approved measures for 30-day mortality of AMI and heart failure patients in that inappropriate avoidance of high risk PCI cases may have a detrimental effect on hospitals' performance on these other measures of cardiovascular outcomes.</p> <p>Describe how could these potential problems be audited: As discussed above, measure implementation will require close attention to data quality. Potential solutions include a) detailed chart audits, b) close attention to variances in case mix and c) review of some or all cases coded as cardiogenic shock or a salvage PCI.</p> <p>Did you audit for these potential problems during testing? No If yes, provide results: N/A</p>
39 (4e)	<p>Testing feasibility Describe what have you learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: As noted previously, testing of this measure has not been performed. However, the NCDR CathPCI in-hospital mortality measure has already been approved by NQF. The implementation of this measure for benchmarking hospital performance demonstrates the feasibility of the proposed 30-day measures with regards to data collection, missing data, and data quality.</p>
CONTACT INFORMATION	
40	<p>Web Page URL for Measure Information Describe where users (implementers) should go for more details on specifications of measures, or assistance in implementing the measure. Web page URL: N/A</p>
41	<p>Measure Intellectual Property Agreement Owner Point of Contact First Name: MI: Last Name: Credentials (MD, MPH, etc.): Organization: Street Address: City: State: ZIP: Email: Telephone: ext:</p>
42	<p>Measure Submission Point of Contact If different than IP Owner Contact First Name: Lein MI: F Last Name: Han Credentials (MD, MPH, etc.): PhD Organization: Centers for Medicare & Medicaid Services (CMS) Street Address: 7500 Security Blvd City: Baltimore State: MD ZIP: 21244-9045 Email: Lein.Han@cms.hhs.gov Telephone: 410-786-0205 ext:</p>
43	<p>Measure Developer Point of Contact If different than IP Owner Contact First Name: Harlan MI: M Last Name: Krumholz Credentials (MD, MPH, etc.): MD Organization: Yale/YNHH Center for Outcomes Research and Evaluation (YNHH-CORE) Street Address: 1 Church Street, Suite 200 City: New Haven State: CT ZIP: 06510-3330 Email: harlan.krumholz@yale.edu Telephone: 203-764-9659 ext:</p>
44	<p>Measure Steward Point of Contact If different than IP Owner Contact <i>Identifies the organization that will take responsibility for updating the measure and assuring it is consistent with the scientific evidence and current coding schema; the steward of the measure may be different than the developer.</i> First Name: Lein MI: F Last Name: Han Credentials (MD, MPH, etc.): PhD Organization: Centers for Medicare & Medicaid Services (CMS) Street Address: 7500 Security Blvd City: Baltimore State: MD ZIP: 21244-9045 Email: Lein.Han@cms.hhs.gov Telephone: 410-786-0205 ext</p>
ADDITIONAL INFORMATION	

<p>45</p>	<p>Workgroup/Expert Panel involved in measure development Workgroup/panel used</p> <p>► If workgroup used, describe the members' role in measure development: The measure developer, Yale/Yale-New Haven Hospital Center for Outcomes Research and Evaluation (YNHH-CORE) obtained expert and stakeholder input on the two measures through two mechanisms. First, the team has held regular conference calls with a Working Group of YNHH-CORE and American College of Cardiology (ACC)/National Cardiovascular Data Registry (NCDR) experts in cardiovascular registries and in the outcomes measure field. Second, YNHH-CORE sought and considered the input of an American College of Cardiology Foundation (ACCF) designated Task Force.</p> <p>► Provide a list of workgroup/panel members' names and organizations:</p> <p>Working Group</p> <p>Ralph Brindis, M.D., M.P.H., F.A.C.C. Regional Senior Advisor for Cardiovascular Disease, Northern California Kaiser Permanente; Clinical Professor of Medicine, UCSF, Oakland, CA; Chief Medical Officer and Chairman, Management Board, National Cardiovascular Data Registry</p> <p>Barbara Christensen, R.N., M.H.A. Senior Director, Registry Services, American College of Cardiology</p> <p>Jeptha Curtis, M.D. Assistant Professor of Medicine, Department of Internal Medicine (Cardiovascular Disease), Yale University</p> <p>Elizabeth Drye, M.D., S.M. Research Project Director, Yale/Yale-New Haven Hospital Center for Outcomes Research and Evaluation</p> <p>Susan Fitzgerald, R.N., M.B.A. Associate Director, Registry Development, American College of Cardiology</p> <p>Lori Geary, M.P.H. Research Project Coordinator, Yale/Yale-New Haven Hospital Center for Outcomes Research and Evaluation</p> <p>Amy Heller, Ph.D., M.P.H. Associate Director, Quality Products, American College of Cardiology</p> <p>Tony Hermann, R.N., M.B.A., C.P.H.Q. Associate Director, CathPCI Registry, American College of Cardiology</p> <p>Kathleen Hewitt, R.N., M.S.N., C.P.H.Q. Associate Vice President, American College of Cardiology</p> <p>Harlan Krumholz, M.D., M. Sc., F.A.C.C. Director, Yale Center for Outcomes Research and Evaluation; Representative, NCDR analytic center; Ex-officio to Task Force</p> <p>Kristi Mitchell, M.P.H. Senior Director, Research, Development and Quality Products, American College of Cardiology</p> <p>Eric Peterson, M.D., M.P.H., F.A.C.C. Professor of Medicine, Duke University; Director, Cardiovascular Outcomes, Duke Clinical Research Institute, Chapel Hill, NC; Member, NCDR Science Oversight Committee/ Representative, NCDR Analytic Center</p> <p>John Rumsfeld, M.D., Ph.D., F.A.C.C. Associate Professor of Medicine, University of Colorado; Clinical Coordinator, VA Ischemic Heart Disease QUE, Denver, CO; Chief Science Officer, National Cardiovascular Data Registry</p>
-----------	--

Lara Slattery, M.H.S.

Director, Quality Services Department - Registries, Products, and Publishing Division, American College of Cardiology

John Spertus, M.D., M.P.H., F.A.C.C.

Director of Cardiovascular Education and Outcomes Research, Mid America Heart Institute, Kansas City, MO; Member, NCDR Science Oversight Committee/Representative, NCDR analytic center; Chair, American College of Cardiology Foundation Task Force on Public Reporting of Hospital-Level Outcomes Measures

Yongfei Wang, M.S.

Senior Research Analyst, Yale/Yale-New Haven Hospital Center for Outcomes Research and Evaluation

William Weintraub, M.D., F.A.C.C.

Chair, CathPCI Registry Steering Committee; Section Chief, Cardiology, Christiana Care Health Services, Inc., Newark DE

Al Woodward, Ph.D., M.B.A.

Director, Research Services, American College of Cardiology

Task Force

Five Task Force members also serve as members of the Working Group, including:

Ralph G. Brindis, M.D., M.P.H., F.A.C.C.

Harlan Krumholz, M.D., M. Sc., F.A.C.C.

Eric Peterson, M.D., M.P.H., F.A.C.C.

John Rumsfeld, M.D., Ph.D., F.A.C.C.

John Spertus, M.D., M.P.H., F.A.C.C.

Other Task Force members are:

John Brush, M.D., F.A.C.C.

Cardiology Consultants LLC, Norfolk, VA; Chair, Quality Strategic Directions Committee

Vincent J. Bufalino, M.D., F.A.C.C.

Midwest Heart Specialists, Naperville, IL; Co-Chair, ACC Advocacy Committee

Gregory Dehmer, M.D., F.A.C.C.

Professor of Medicine, Texas A&M College of Medicine, Temple, TX;

Representative, The Society for Cardiovascular Angiography and Interventions

James Dove, M.D., F.A.C.C.

President, American College of Cardiology

President Emeritus, Prairie Cardiovascular Consultants, Ltd., Springfield, IL; President, ACC/ACCF Board of Trustees

Stephen C. Hammill, M.D., F.H.R.S.

Professor of Medicine, Mayo Clinic College of Medicine, Rochester, MN; Representative, Heart Rhythm Society

Frank E Harrell Jr., PhD

Professor of Biostatistics; Department Chair, Vanderbilt University School of Medicine- Department of Biostatistics, Nashville, TN

Barry K. Lewis, D.O., F.A.C.C.

Consultants in Cardiology, P.C., Farmington Hills, MI; Member, Advocacy Committee

	<p>William R. Lewis, M.D., F.A.C.C. Metro Health Medical Center, Cleveland, OH; ACC Ohio Chapter Governor/ACC Board of Governors</p> <p>Fred Masoudi, M.D., M.S.P.H., F.A.C.C. Denver Health Medical Center, Denver, CO; Chair, ACC/AHA Task Force on Performance Measures</p> <p>Andrea M. Russo, M.D. F.A.C.C. University of Pennsylvania Health System, Philadelphia, PA; Representative, Heart Rhythm Society</p> <p>Bonnie H. Weiner, M.D., F.S.C.A.I., F.A.C.C. Professor of Medicine; Interim Chair Cardiovascular Medicine, St. Vincent Hospital at Worcester Medical Center, Worcester, MA; Representative, The Society for Cardiovascular Angiography and Interventions</p> <p>Stuart Winston, D.O., F.A.C.C. Michigan Heart, P. C., Ann Arbor, MI; ACC Michigan Chapter Governor/ACC Board of Governors</p>
46	<p>Measure Developer/Steward Updates and Ongoing Maintenance</p> <p>Year the measure was first released: N/A</p> <p>Month and Year of most recent revision: N/A</p> <p>What is the frequency for review/update of this measure? N/A</p> <p>When is the next scheduled review/update for this measure? N/A</p>
47	<p>Copyright statement/disclaimers: N/A</p>
48	<p>Additional Information: Hospital 30-day Percutaneous Coronary Intervention Mortality Measures Methodology Report (attached)</p>
49	<p>I have checked that the submission is complete and any blank fields indicate that no information is provided. <input checked="" type="checkbox"/></p>
50	<p>Date of Submission (MM/DD/YY): 11-21-08</p>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

The measure information you submit will be shared with NQF’s Steering Committees and Technical Advisory Panels to evaluate measures against the NQF criteria of importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Four conditions (as indicated below) must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, *importance to measure and report*, in order to be evaluated against the remaining criteria. References to the specific measure evaluation criteria are provided in parentheses following the item numbers. Please refer to the *Measure Evaluation Criteria* for more information at www.qualityforum.org under Core Documents. Additional guidance is being developed and when available will be posted on the NQF website.

Use the tab or arrow (↓→) keys to move the cursor to the next field (or back ←↑). There are three types of response fields:

- drop-down menus - select one response;
- check boxes - check as many as apply; and
- text fields - you can copy and paste text into these fields or enter text; these fields are not limited in size, but in most cases, we ask that you summarize the requested information.

Please note that URL hyperlinks do not work in the form; you will need to type them into your web browser.

Be sure to answer all questions. Fields that are left blank will be interpreted as no or none. **Information must be provided in this form.** Attachments are not allowed except to provide additional detail or source documents for information that is summarized in this form. If you have important information that is not addressed by the questions, they can be entered into item #46 near the end of the form.

For questions about this form, please contact the NQF Project Director listed in the corresponding call for measures.

CONDITIONS FOR CONSIDERATION BY NQF	
	Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards.
A (A)	Public domain or Measure Steward Agreement signed: Public domain - Agreement not required (If no, do not submit) <i>Template for the Measure Steward Agreement is available at www.qualityforum.org under Core Documents.</i>
B (B)	Measure steward/maintenance: Is there an identified responsible entity and process to maintain and update the measure on a schedule commensurate with clinical innovation, but at least every 3 years? Yes, information provided in contact section (If no, do not submit)
C (C)	Intended use: Does the intended use of the measure include BOTH public reporting AND quality improvement? Yes (If no, do not submit)
D (D)	Fully developed and tested: Is the measure fully developed AND tested? Yes, fully developed and tested (If not tested and no plans for testing within 24 months, do not submit)

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

	(for NQF staff use) NQF Review #: HOE-019-08	NQF Project: Hospital Outcomes and Efficiency
MEASURE SPECIFICATIONS & DESCRIPTIVE INFORMATION		
1	Information current as of (date- MM/DD/YY):	
2	Title of Measure: Survival Predictor for CABG Surgery	
3	Brief description of measure ¹ : A reliability adjusted measure of CABG surgical performance that optimally combines two important domains: CABG hospital volume and CABG operative mortality, to provide predictions on CABG survival rates for hospitals. This measure is calculated based on data from administrative claims information.	
4 (2a)	<p>Numerator Statement: Note: Because of the type of modeling done for this Survival Predictor--the information is not readily split into Numerator/ Denominator statements. Thus, we describe the two domains and their coding and data needs in this section.</p> <p>The formula for calculating the survival predictor has two components, one is a volume predicted mortality rate, and the second is an observed mortality rate. The volume predicted mortality rate reflects the hospitals experience performing CABG surgeries (thus, it includes all CABG surgeries) and uses mortality for all hospitals at that specific volume to create the volume predicted mortality. The input data from the hospitals for this domain is a volume count of all CABGs performed in the hospital.</p> <p>The second domain is the observed mortality, for this domain the population is narrowed to a homogenous group of isolated CABG cases, the data needed for this domain is the number of observed deaths occurring for isolated CABG cases, within the inpatient setting.</p> <p>Note: All data is available in administrative claims information. In the case of Leapfrog's implementation hospitals are asked to submit aggregated information from their claims data. No personal health information is submitted to Leapfrog. Other users of the measure may have direct access to administrative data.</p> <p>Time Window: 12 months</p> <p>Numerator Details (Definitions, codes with description): For the volume predicted mortality, hospitals count the number of CABG cases using the following codes: (NQF Endorsed -#0124a,c)</p> <p>ICD-9-CM Procedure</p> <ul style="list-style-type: none"> ■ 36.10 Aortocoronary bypass for heart revascularization,NOS ■ 36.11 Aortocoronary bypass of one coronary artery ■ 36.12 Aortocoronary bypass of two coronary arteries ■ 36.13 Aortocoronary bypass of three coronary arteries ■ 36.15 Single internal mammary-coronary artery bypass ■ 36.16 Double internal mammary-coronary artery bypass ■ 36.19 Other bypass anastomosis for heart revascularization <p>See calculation worksheet for details on how volume-predicted mortality is used in the model.</p> <p>For the observed mortality domain, the hospital submits the observed deaths for isolated CABG cases using</p>	

¹ Example of measure description: Percentage of adult patients with diabetes aged 18-75 years receiving one or more A1c test(s) per year.
NQF Measure Submission Form, V3.1

	<p>the following codes (NQF endorsed #0124a):</p> <ul style="list-style-type: none"> 36.10 Aortocoronary bypass for heart revascularization,NOS ■ 36.11 Aortocoronary bypass of one coronary artery ■ 36.12 Aortocoronary bypass of two coronary arteries ■ 36.13 Aortocoronary bypass of three coronary arteries ■ 36.15 Single internal mammary-coronary artery bypass ■ 36.16 Double internal mammary-coronary artery bypass ■ 36.19 Other bypass anastomosis for heart revascularization <p>And, from that group of CABG cases, hospitals exclude cases with a valve procedure; codes for the exclusion are:</p> <p>ICD-9-CM Procedure Codes:</p> <ul style="list-style-type: none"> 35.10 Open heart valvuloplasty without replacement, unspecified valve ■ 35.11 Open heart valvuloplasty of aortic valve without replacement ■ 35.12 Open heart valvuloplasty of mitral valve without replacement ■ 35.13 Open heart valvuloplasty of pulmonary valve without replacement ■ 35.14 Open heart valvuloplasty of tricuspid valve without replacement ■ 35.20 Replacement of unspecified heart valve ■ 35.21 Replacement of aortic valve with tissue graft ■ 35.22 Other replacement of aortic valve ■ 35.23 Replacement of mitral valve with tissue graft ■ 35.24 Other replacement of mitral valve ■ 35.25 Replacement of pulmonary valve with tissue graft ■ 35.26 Other replacement of pulmonary valve ■ 35.27 Replacement of tricuspid valve with tissue graft ■ 35.28 Other replacement of tricuspid valve <p>Thus, the observed mortality is based on the volume count of isolated CABGs and an actual count of deaths occurring for that subset of CABG cases.</p> <p>See Calculation Worksheet for how the two domains are used to create the Survival Predictor.</p>
<p>5 (2a)</p>	<p>Denominator Statement: See numerator section for all data needed, and codes</p> <p>Time Window:</p> <p>Denominator Details (Definitions, codes with description):</p>
<p>6 (2a, 2d)</p>	<p>Denominator Exclusions: These exclusions are for the observed mortality domain--it excludes all cases with concomitant valve replacement or repair (essentially the definition of isolated CABG.)</p> <p>Denominator Exclusion Details (Definitions, codes with description): (These are valve codes from the endorsed CMS measure (# 0124b))</p> <p>ICD-9-CM Procedure Codes: 35.10-35.29</p> <ul style="list-style-type: none"> 35.10 Open heart valvuloplasty without replacement, unspecified valve ■ 35.11 Open heart valvuloplasty of aortic valve without replacement ■ 35.12 Open heart valvuloplasty of mitral valve without replacement ■ 35.13 Open heart valvuloplasty of pulmonary valve without replacement ■ 35.14 Open heart valvuloplasty of tricuspid valve without replacement ■ 35.20 Replacement of unspecified heart valve ■ 35.21 Replacement of aortic valve with tissue graft ■ 35.22 Other replacement of aortic valve ■ 35.23 Replacement of mitral valve with tissue graft ■ 35.24 Other replacement of mitral valve

	<ul style="list-style-type: none"> ■ 35.25 Replacement of pulmonary valve with tissue graft ■ 35.26 Other replacement of pulmonary valve ■ 35.27 Replacement of tricuspid valve with tissue graft ■ 35.28 Other replacement of tricuspid valve 		
7 (2a, 2h)	<p>Stratification Do the measure specifications require the results to be stratified? No</p> <p>► If "other" describe:</p> <p>Identification of stratification variable(s):</p> <p>Stratification Details (Definitions, codes with description):</p>		
8 (2a, 2e)	<p>Risk Adjustment Does the measure require risk adjustment to account for differences in patient severity before the onset of care? No ► If yes, (select one)</p> <p>► Is there a separate proprietary owner of the risk model? No</p> <p>Identify Risk Adjustment Variables: See section 28 for rationale and support for not risk adjusting this measure. Measure was tested against risk adjusted mortality--details on that provided in Section 26.</p> <p>Detailed risk model: attached <input type="checkbox"/> OR Web page URL:</p>		
9 (2a)	<p>Type of Score: Rate/proportion Calculation Algorithm: attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p>Interpretation of Score (<i>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</i>)</p> <p>Better quality = Score within a defined interval ► If "Other", please describe:</p>		
10 (2a, 4a, 4b)	<p>Identify the required data elements(e.g., primary diagnosis, lab values, vital signs): procedure codes</p> <p>Data dictionary/code table attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p>Data Quality (2a) <i>Check all that apply</i></p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> Data are captured from an authoritative/accurate source (e.g., lab values from laboratory personnel) <input checked="" type="checkbox"/> Data are coded using recognized data standards <input checked="" type="checkbox"/> Method of capturing data electronically fits the workflow of the authoritative source <input type="checkbox"/> Data are available in EHRs <input checked="" type="checkbox"/> Data are auditable 		
11 (2a, 4b)	<p>Data Source and Data Collection Methods <i>Identifies the data source(s) necessary to implement the measure specifications. Check all that apply</i></p> <table style="width: 100%; border: none;"> <tr> <td style="width: 50%; vertical-align: top;"> <ul style="list-style-type: none"> <input type="checkbox"/> Electronic Health/Medical Record <input type="checkbox"/> Electronic Clinical Database, Name: <input type="checkbox"/> Electronic Clinical Registry, Name: <input checked="" type="checkbox"/> Electronic Claims <input type="checkbox"/> Electronic Pharmacy data <input type="checkbox"/> Electronic Lab data <input type="checkbox"/> Electronic source - other, Describe: </td> <td style="width: 50%; vertical-align: top;"> <ul style="list-style-type: none"> <input type="checkbox"/> Paper Medical Record <input type="checkbox"/> Standardized clinical instrument, Name: <input type="checkbox"/> Standardized patient survey, Name: <input type="checkbox"/> Standardized clinician survey, Name: <input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period. </td> </tr> </table> <p style="text-align: right;">Instrument/survey attached <input checked="" type="checkbox"/> OR Web page URL:</p>	<ul style="list-style-type: none"> <input type="checkbox"/> Electronic Health/Medical Record <input type="checkbox"/> Electronic Clinical Database, Name: <input type="checkbox"/> Electronic Clinical Registry, Name: <input checked="" type="checkbox"/> Electronic Claims <input type="checkbox"/> Electronic Pharmacy data <input type="checkbox"/> Electronic Lab data <input type="checkbox"/> Electronic source - other, Describe: 	<ul style="list-style-type: none"> <input type="checkbox"/> Paper Medical Record <input type="checkbox"/> Standardized clinical instrument, Name: <input type="checkbox"/> Standardized patient survey, Name: <input type="checkbox"/> Standardized clinician survey, Name: <input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.
<ul style="list-style-type: none"> <input type="checkbox"/> Electronic Health/Medical Record <input type="checkbox"/> Electronic Clinical Database, Name: <input type="checkbox"/> Electronic Clinical Registry, Name: <input checked="" type="checkbox"/> Electronic Claims <input type="checkbox"/> Electronic Pharmacy data <input type="checkbox"/> Electronic Lab data <input type="checkbox"/> Electronic source - other, Describe: 	<ul style="list-style-type: none"> <input type="checkbox"/> Paper Medical Record <input type="checkbox"/> Standardized clinical instrument, Name: <input type="checkbox"/> Standardized patient survey, Name: <input type="checkbox"/> Standardized clinician survey, Name: <input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period. 		
12 (2a)	<p>Sampling <i>If measure is based on a sample, provide instructions and guidance on sample size.</i></p> <p>Minimum sample size: h1</p> <p>Instructions:</p>		
13 (2a)	<p>Type of Measure: Outcome ► If "Other", please describe:</p> <p>► If part of a composite or paired with another measure, please identify composite or paired measure</p> <p>While the measure to two types of information components, the results are not a composite as is defined by NQF, but rather a reliability adjusted measure of survival.</p> <p>Volume is used to create a volume predicted mortality for the hospital--this component of the measure is used to create greater reliability for low-volume hospitals. In the modeling for this measure, the volume</p>		

	<p>predicted mortality and the observed mortality are weighted. In the model, lower volume hospitals have a higher weight on the volume predicted mortality versus the observed mortality. The opposite is true for high volume hospitals, which have a higher weight on the observed mortality. This methodology results in a reliability adjusted survival predictor.</p>																		
14 (2a)	<p>Unit of Measurement/Analysis (Who or what is being measured) Check all that apply.</p> <table border="0"> <tr> <td><input type="checkbox"/> Can be measured at all levels</td> <td><input type="checkbox"/> Integrated delivery system</td> </tr> <tr> <td><input type="checkbox"/> Individual clinician (e.g., physician, nurse)</td> <td><input checked="" type="checkbox"/> Health plan</td> </tr> <tr> <td><input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)</td> <td><input checked="" type="checkbox"/> Community/Population</td> </tr> <tr> <td><input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)</td> <td><input type="checkbox"/> Other (Please describe):</td> </tr> </table>	<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system	<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input checked="" type="checkbox"/> Health plan	<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input checked="" type="checkbox"/> Community/Population	<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other (Please describe):										
<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system																		
<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input checked="" type="checkbox"/> Health plan																		
<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input checked="" type="checkbox"/> Community/Population																		
<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other (Please describe):																		
15 (2a)	<p>Applicable Care Settings Check all that apply</p> <table border="0"> <tr> <td><input type="checkbox"/> Can be used in all healthcare settings</td> <td><input type="checkbox"/> Hospice</td> </tr> <tr> <td><input type="checkbox"/> Ambulatory Care (office/clinic)</td> <td><input checked="" type="checkbox"/> Hospital</td> </tr> <tr> <td><input type="checkbox"/> Behavioral Healthcare</td> <td><input type="checkbox"/> Long term acute care hospital</td> </tr> <tr> <td><input type="checkbox"/> Community Healthcare</td> <td><input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)</td> </tr> <tr> <td><input type="checkbox"/> Dialysis Facility</td> <td><input type="checkbox"/> Prescription Drug Plan</td> </tr> <tr> <td><input type="checkbox"/> Emergency Department</td> <td><input type="checkbox"/> Rehabilitation Facility</td> </tr> <tr> <td><input type="checkbox"/> EMS emergency medical services</td> <td><input type="checkbox"/> Substance Use Treatment Program/Center</td> </tr> <tr> <td><input type="checkbox"/> Health Plan</td> <td><input type="checkbox"/> Other (Please describe):</td> </tr> <tr> <td><input type="checkbox"/> Home Health</td> <td></td> </tr> </table>	<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice	<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital	<input type="checkbox"/> Behavioral Healthcare	<input type="checkbox"/> Long term acute care hospital	<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)	<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan	<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility	<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center	<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other (Please describe):	<input type="checkbox"/> Home Health	
<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice																		
<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital																		
<input type="checkbox"/> Behavioral Healthcare	<input type="checkbox"/> Long term acute care hospital																		
<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)																		
<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan																		
<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility																		
<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center																		
<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other (Please describe):																		
<input type="checkbox"/> Home Health																			
<p>IMPORTANCE TO MEASURE AND REPORT</p>																			
<p>Note: This is a threshold criterion. If a measure is not judged to be sufficiently important to measure and report, it will not be evaluated against the remaining criteria.</p>																			
16 (1a)	<p>Is measure related to a National Priority Partners priority area? Safety reliability (for NQF staff use) Does measure address a <u>specific</u> NPP goal? (www.qualityforum.org/about/NPP/):</p>																		
17 (1a)	<p>Does the measure address a high impact aspect of healthcare patient/societal consequences of poor quality</p> <p>Summary of Evidence: This measure addresses mortality in a high risk procedure (CABG) and is an outcome measure which is of interest to both consumers and purchasers.</p> <p>In 2006, there were 444,000 CABG procedures done in US hospitals [1] with more than \$52 million was spent on CABG surgeries, this is 5.6% of the national bill for hospitals; it also accounts for 1.2 million hospital stays. It is number 1 in the top 20 most expensive conditions treated in US hospitals. [2]</p> <p>Mortality in US hospitals varies for CABG surgeries--there are documented differences between high and low performing hospitals [4]. Higher volumes are associated with better outcomes including lower mortality.</p> <p>In addition to addressing a high risk, high cost procedure, this measure improves upon the technology of surgical mortality measurement. It overcomes three problems with existing CABG mortality measures: 1) Mortality rates are often too "noisy" to reflect hospital quality with surgery (particularly among lower volume hospitals), 2) volume alone is a weak proxy for most procedures, and 3) when both volume and mortality are reported as separate indicators it is difficult to understand which measure is more important. [1] Given the large number of CABG procedures performed annually in the United States, and that this measure specifically addresses hospitals which perform elective procedures, consumers and purchasers would benefit from information that is more reliable in the prediction of future mortality for both selection and selective referral. In addition, this measure can be applied to the nation, states, or regions.</p> <p>Birkmeyer and Dimick (2009)[4] show that differences in mortality can be predicted using a reliability</p>																		

	<p>adjusted mortality rate (a weighted combination of volume and mortality) which is particularly relevant for selective-referral or public reporting contexts. They reduce the effects of random chance (statistical noise) and as a result with CABG, for example, more than half of the observed variation can be attributed to statistical noise. When they sorted hospitals simply on actual (risk-adjusted) mortality, rates varied from 1.4% to 11.0% across hospital quintiles (Figure 1 in White Paper [1]). After they adjusted for reliability, however, the mortality rates varied considerably less, from 3.3% to 6.3%. Although the almost twofold variation in mortality still suggests ample opportunity for quality improvement, these data underscore the importance of accounting for chance in understanding variation in hospital outcomes.</p> <p>Citations² for Evidence:</p> <p>[1] DeFrances, C.J., Lucas, CA, Bule, VC., Golosinskiy, A. 2006 National Hospital Discharge Survey, National health statistics reports, no. 5. Hyattsville, MD: National Center for Health Statistics. 2008. Accessed on 12/17/08 at http://www.cdc.gov/nchs/data/nhsr/nhsr005.pdf</p> <p>[2] The National Hospital Bill: The Most Expensive Conditions by Payer, 2006. Statistical Brief #59. File accessed on March 16, 2009, at: http://www.hcup-us.ahrq.gov/reports/statbriefs/sb59.jsp Produced by AHRO, Center for Delivery, Organization, and Markets, Healthcare Cost and Utilization Project, Nationwide Inpatient Sample, 2006.</p> <p>[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf</p> <p>[4] Birkmeyer, J.D., and Dimick, J.B. (2009) Understanding and reducing variation in surgical mortality. <i>Annu. Rev. Med.</i> 2009. 60:405-15.</p>
<p>18 (1b)</p>	<p>Opportunity for Improvement <i>Provide evidence that demonstrates considerable variation, or overall poor performance, across providers.</i></p> <p>Summary of Evidence: In 2002, a systematic review of the literature on the volume-outcome relationship found that there was a significant relationship between hospital volume and outcomes for CABG surgery. While this relationship was not as robust in CABG surgery as it was in some other surgical procedures (esophagectomy, pancreatectomy), it was present. [5]</p> <p>In 2005, Epstein, Rathore, Krumholz, et al., [6] found that moving patients for CABG procedures from low volume hospitals to high volume hospitals meeting Leapfrog's standard would save lives--the mortality odds ratio for low volume hospitals was 1.16 (95% C.I. 1.10-1.24).</p> <p>Given the findings related to volume of procedures, Silber et al., [7] explored the relative contribution of complication rates and failure to rescue rates to mortality and found that complication rates were more likely influenced by patient factors while failure to rescue rates of those with complications was more related to hospital factors. Thus, it may be that higher volume hospitals are better at rescuing patients with complications. Silbers finding, in conjunction with the volume information, suggests lower volume hospitals with worse mortality rates could in fact address this through better care following the procedure, thereby reducing their overall rate. Unfortunately, most low volume hospitals in the United States do not have information on their CABG mortality rate compared to other hospitals. When they are given this information, there is a good chance for improvement. Birkmeyer and Dimick [4] report that in northern New England, mortality associated with CABG fell by >25% when hospitals and surgeons were given feedback on their mortality data.</p> <p>Note: Birkmeyer and Dimmick [4] indicate it is also likely that some lower volume hospitals would also have lower mortality rates.</p>

² Citations can include, but are not limited to journal articles, reports, web pages (URLs).
NQF Measure Submission Form, V3.1

	<p>Citations for Evidence: [5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. <i>Annals of Internal Medicine</i>, Sept 1;137(6):511-20</p> <p>[6] Epstein, A.J., Rathore, S.S., Krumholz, H.M., and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>[7] Silber, J.H., Rosenbaum, P.R., Trudeau, M.E., et al. 2005. Changes in prognosis after the first postoperative complication. <i>Medical Care</i>, 43:122-31.</p>
<p>19 (1b)</p>	<p>Disparities <i>Provide evidence that demonstrates disparity in care/outcomes related to the measure focus among populations.</i></p> <p>Summary of Evidence: It is more likely that minorities will be treated at a low volume facility, and as a result are likely to be impacted by higher mortality rates. In an analysis of the National Inpatient Sample, Epstein, Rathore and Krumholz (2005)[6] found that a greater proportion of patients treated in low volume hospitals were non-white, while a lower proportion of non-white patients presented as "elective" admissions or patients received in transfer as compared to patients in high volume hospitals.</p> <p>The impact of socioeconomic status in relation to surgical mortality was studied by Birkmeyer (N.J.O.) et al., (2008). They found that for CABG surgery the odds ratio for lower SES was 1.14 (95% C.I.:1.09-1.19). The disparities in surgical outcomes were attributed to differences in hospitals where low and high SES patients sought surgical treatment. [8]</p> <p>In the survival predictor, the denominator and numerator are restricted to elective procedures, therefore, it is anticipated there may be a smaller non-white, and low SES population in the denominator and numerator.</p> <p>Citations for evidence:</p> <p>[4, p. 3-5]</p> <p>[8] Birkmeyer, N.J.O., Gu, N., Baser, O., Morris, A.M., and Birkmeyer, J.D. (2008). <i>Medical Care</i>; 46:893-899.</p>
<p>20 (1c)</p>	<p>If measuring an Outcome Describe relevance to the national health goal/priority, condition, population, and/or care being addressed: A CABG procedure is a high risk procedure, and a very expensive procedure, and only limited information is available nationally on the risk of mortality associated with the procedure. Other entities with clinical information are not publicly reporting mortality rates of CABG procedures by hospital provider. A few states (NY, CA, NJ, PA, MA) are reporting on the CABG surgeries performed in their state, however, these reports have significant delays in reporting. Prominent private registries are not producing public information, e.g., Society for Thoracic Surgery. This measure is designed to give feedback to hospitals across the country as well as to provide information for decision-making by consumers and purchasers.</p> <p>Mortality in US hospitals varies for CABG surgeries--there are documented differences between high and low performing hospitals [4]. Higher volumes are associated with better outcomes including lower mortality. In addition to being a high risk surgery, this surgery is one of the high cost procedures, both by single event and by total in this country. In 2006, more than \$52 million was spent on CABG surgeries, this is 5.6% of the national bill for hospitals; it also accounts for 1.2 million hospital stays. It is number 1 in the top 20 most expensive conditions treated in US hospitals. [2]</p> <p>This measure is highly relevant to both consumers and purchasers, given its high cost both in terms of lives lost and dollars spent. National purchasers are interested in comparative information on hospitals nationwide. Pauly (1996) in a study of purchaser interests in hospital performance reporting found that mortality ratings were more important to purchasers than were morbidity or complications. [9] Health plans are interested in contracting with centers of excellence, which can be identified through the results</p>

of survival predictor in combination with other information on cost and quality. Consumers have shown their interest in CABG mortality by requesting reports from the state of Pennsylvania [10]; an earlier study by IOM (Lohr, Donaldson and Walker 1991) found that consumers were interested in hospital mortality rates, but did not perceive this information to be available.[11] Hibbard and Jewett found that consumers were more interested in "undesirable events" (such as mortality, complications, infections) than in "desirable events." [12]

[9] Pauly, M.V., Brailer, D.J.Kroch, E., and Even-Shoshan, O. Measuring Hospital Outcomes from a Buyer's Perspective. American Journal of Medical Quality, 11(8): Fall 1996.

[10] Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.

[11] Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. Quality Review Bulletin 17:242-53.

[12] Hibbard, J.H. and Jewett, J. (1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? Medical Care Research and Review., Vol 53(1): 28-47.

If not measuring an outcome, provide evidence supporting this measure topic and grade the strength of the evidence

Summarize the evidence (including citations to source) supporting the focus of the measure as follows:

- Intermediate outcome - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
- Process - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
- Structure - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
- Patient experience - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
- Access - evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
- Efficiency- demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

Type of Evidence Check all that apply

- | | |
|--|---|
| <input type="checkbox"/> Evidence-based guideline | <input checked="" type="checkbox"/> Quantitative research studies |
| <input checked="" type="checkbox"/> Meta-analysis | <input type="checkbox"/> Qualitative research studies |
| <input checked="" type="checkbox"/> Systematic synthesis of research | <input type="checkbox"/> Other (Please describe): |

Overall Grade for Strength of the Evidence³ (Use the USPSTF system, or if different, also describe how it relates to the USPSTF system): **Moderate**

Summary of Evidence (provide guideline information below): **Over 100 articles published related to volume and outcome relationship, with some inconsistency in results. Systematic review of the literature conducted in 2002. No review since that time.**

³The strength of the body of evidence for the specific measure focus should be systematically assessed and rated, e.g., USPSTF grading system www.ahrq.gov/clinic/uspstmeth.htm: **A** - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. **B** - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. **C** - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. **D** - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. **I** - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

	<p>Citations for Evidence:</p> <p>[5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. <i>Annals of Internal Medicine</i>, Sept 1;137(6):511-20.</p> <p>[14] Birkmeyer, J.D., Dimick, J.B., Staiger, D.O. (2006) Operative mortality and procedure volume as predictors of subsequent hospital performance. <i>Ann Surg.</i> 243:411-417.</p> <p>[15] Dimick, JB, Welch HG, Birkmeyer JD. (2004) Surgical mortality as an indicator of hospital quality: The problem with small sample size. <i>JAMA</i>, 292:847-851.</p> <p>[4] Birkmeyer, JD., and Dimick, JB. (2009) Understanding and reducing variation in surgical mortality. <i>Annu. Rev. Med.</i> 60:405-15.</p> <p>[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. <i>Medical Care</i>, 47(2): 226-233.</p> <p>[6] Andrew J Epstein, Saif S Rathore, Harlan M Krumholz and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>[17] Edward L. Hannan, PhD; Chuntao Wu, PhD; Thomas J. Ryan, MD; Edward Bennett, MD; Alfred T. Culliford, MD; Jeffrey P. Gold, MD; Alan Hartman, MD; O. Wayne Isom, MD; Robert H. Jones, MD; Barbara McNeil, MD, PhD; Eric A. Rose, MD; Valavanur A. Subramanian, MD. Do Hospitals and Surgeons With Higher Coronary Artery Bypass Graft Surgery Volumes Still Have Lower Risk-Adjusted Mortality Rates? <i>Circulation</i>. 2003;108:795-801.</p> <p>[18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. <i>N Engl J Med</i>. 1979;301:1364-9.</p>
<p>21 (1c)</p>	<p>Clinical Practice Guideline <i>Cite the guideline reference; quote the specific guideline recommendation related to the measure and the guideline author's assessment of the strength of the evidence; and summarize the rationale for using this guideline over others.</i></p> <p>Guideline Citation:</p> <p>Specific guideline recommendation:</p> <p>Guideline author's rating of strength of evidence (If different from USPSTF, also describe it and how it relates to USPSTF):</p> <p>Rationale for using this guideline over others:</p>
<p>22 (1c)</p>	<p>Controversy/Contradictory Evidence <i>Summarize any areas of controversy, contradictory evidence, or contradictory guidelines and provide citations.</i></p> <p>Summary: There are three areas of possible contention with this measure--</p> <p>1) The volume-outcome relationship has been questioned for some procedures [6, 17, 19]</p> <p>Peterson et al., [19] questioned the volume outcomes relationship for CABG surgery, and found only modest associations for volume and outcome for CABG. Those with high volume had mortality rate of 2.5% while low volume hospitals rate was 3.2%. They suggest using past mortality rate to select hospitals. (The survival predictor uses both volume and mortality to predict survival in the next year.)</p> <p>More than 100 studies have demonstrated better results at high-volume hospitals with cardiovascular surgery, major cancer resections, and other high-risk procedures.[18, 20]</p> <p>All studies listed here were done to determine whether there was a volume-outcome relationship for</p>

	<p>hospitals performing CABG procedures--they all documented that there were differences between low volume hospitals and high volume hospitals in mortality--with high volume hospitals having less mortality. In the case of CABG surgeries the variance between high and low was more modest than some other high risk procedures.</p> <p>2) That outcome measures must be risk-adjusted unless there is evidence to show it is not needed (NQF). The survival predictor measure predicts better than volume or mortality alone, and is as good a predictor as risk-adjusted mortality. When testing the unadjusted survival predictor against risk-adjusted mortality there was a (.96) correlation. [4] See Section 28 of this form for details.</p> <p>3) The weighting of input measures into composites. Existing approaches rely on overly simplistic approaches. Among these, assigning equal weight to all measures (i.e., the all or none approach) and relying on expert opinion are the most common. The survival predictor relies on empiric methods for weighting the input measures.</p> <p>Citations:</p> <p>[6] Andrew J Epstein, Saif S Rathore, Harlan M Krumholz and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. BMC Health Services Research, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>[16]Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. Medical Care, 47(2): p. 232.</p> <p>[17] Edward L. Hannan, PhD; Chuntao Wu, PhD; Thomas J. Ryan, MD; Edward Bennett, MD; Alfred T. Culliford, MD; Jeffrey P. Gold, MD; Alan Hartman, MD; O. Wayne Isom, MD; Robert H. Jones, MD; Barbara McNeil, MD, PhD; Eric A. Rose, MD; Valavanur A. Subramanian, MD. Do Hospitals and Surgeons With Higher Coronary Artery Bypass Graft Surgery Volumes Still Have Lower Risk-Adjusted Mortality Rates? Circulation. 2003;108:795-801.</p> <p>[18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. N Engl J Med. 1979;301:1364-9.</p> <p>[19] Eric D. Peterson, MD, MPH; Laura P. Coombs, PhD; Elizabeth R. DeLong, PhD; Constance K. Haan, MD; T. Bruce Ferguson, MD. Procedural Volume as a Marker of Quality for CABG Surgery. JAMA. 2004;291:195-201.</p> <p>[20] Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. JAMA. 1998;280:1747-51.)</p>
<p>23 (1)</p>	<p>Briefly describe how this measure (as specified) will facilitate significant gains in healthcare quality related to the specific priority goals and quality problems identified above: This measure of predicted survival improves upon the reliability of mortality results for high risk surgical procedures, such as CABG. For the first time, this measure produces reliable mortality/survivability information on smaller volume hospitals, as well as high volume hospitals. Hospitals across the country will have information available through voluntary public reporting.</p>
<p>SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES</p>	
<p>Note: Testing and results should be summarized in this form. However, additional detail and reports may be submitted as supplemental information or provided as a web page URL. If a measure has not been tested, it is only potentially eligible for time-limited endorsement.</p>	
<p>24</p>	<p>Supplemental Testing Information: attached <input checked="" type="checkbox"/> OR Web page URL:</p>
<p>25 (2b)</p>	<p>Reliability Testing</p> <p>Data/sample: Data was a 100% sample from the Medicare Analysis Provider and Review (MEDPAR) files for 2000-2003, these files contain 100% of Medicare hospitalizations for years specified. MEDPAR files, which contain</p>

hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.

Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the coronary artery bypass grafting surgery. We excluded small patient subgroups with much higher baseline risks, including those with procedure codes indicating that other operations were simultaneously performed (e.g., coronary artery bypass and valve surgery) or were performed for emergent indications. [3]

Note: Needleman, Buerhaus, et al., (2003) concluded after applying operational tests on Medicare data for adverse outcomes and all-patient hospital data from 11 states, that Medicare data could be used to assess quality in hospitals.[20] Given the lack of a national all-patient database, MEDPAR data was used in development and testing of the models.

Analytic Method:

Model Development

We used an empirical Bayes approach to combine mortality rates with information on hospital volume at each hospital. In traditional empirical Bayes methods, a point estimate (e.g., mortality rate observed at a hospital) is adjusted for reliability by shrinking it towards the overall mean (e.g., overall mortality rate in the population) [21,22]. We modified this traditional approach by shrinking the observed mortality rate back toward the mortality rate expected given the volume at that hospital—we refer to this as the “volume-predicted mortality” (See attached White Paper TECHNICAL APPENDIX for the mathematical details of this method). With this approach, the observed mortality rate is weighted according to how reliably it is estimated, with the remaining weight placed on the information regarding hospital volume. Because this method includes observed data to the extent that it is useful, and only relies on the proxy measure to the extent necessary, it ensures an optimal combination of these two quality domains. [3]

The two inputs to the survival predictor measure are mortality rates and procedure volume for each of the six included operations. Procedure-specific mortality rates were calculated for all hospitals over a 2-year period (2000-01) and this was used as the first input. Hospital volume was calculated as the number of Medicare cases performed during the same time period. For each operation, the relationship between hospital volume and risk-adjusted mortality was modeled using linear regression. (Details of the risk-adjustment strategy will be discussed below.) After testing the fit of several transformations, hospital volume was modeled as the natural log of the continuous volume variable, which is the same approach used in our previous work [23]. Using this regression model, we estimated the volume-predicted mortality, the second input to the survival predictor measure.

We then used the empirical Bayes approach to create an optimal combination of these two inputs. This survival predictor measure theoretically provides the best estimate of a hospital's true mortality rate, taking into account the both available inputs [21,22]. The combined survival predictor measure was calculated as follows: mortality prediction = (weight)*(observed mortality) + (1-weight)*(volume-predicted mortality). The weight placed on the point estimate of mortality is the reliability, or ratio of signal to signal plus noise, calculated as follows: weight = variation among hospitals/(variation among hospitals + variation within hospitals). The variation among hospitals was calculated as the variance in observed mortality rates for the hospitals included in the sample. The variation within hospitals was calculated as the standard error of the mortality rate at each hospital. With this method, more weight is placed on the observed mortality rate when a hospital has a high number of cases because it is estimated with more reliability; less weight is placed on the observed mortality rate when a hospital performs a low number of cases because of its lower reliability.

A calculation worksheet is attached.

	<p>Testing Results: Hospital caseloads and the weights applied to each input to the survival predictor measure varied for each procedure studied (see Table 1 in white paper [3]). For coronary artery bypass, a procedure with relatively high hospital caseloads, the weight applied to the volume input was 54%. ([3]-Table 1). For hospitals with higher volumes more weight was placed on the observed mortality.</p> <p>The survival predictor (mortality) measure explained a large proportion of non-random, hospital-level variation in risk-adjusted mortality rates (see Table 2 in White Paper [3]). For coronary artery bypass, the survival predictor explained 61% of the hospital level variation in mortality rates; this compares to 46% for observed mortality and 9% for volume of CABG surgeries. Measures with low reliability or correlation explain little variation. The correlation between the survival predictor and risk-adjusted mortality was (.96) ([16] p. 232), and the amount of variation explained was 61% [3]. This is a more than adequate level of reliability.</p> <p>Citations:</p> <p>[3] p. 19 (Table 2)</p> <p>[21] Morris CN. Parametric Empirical Bayes Inference: Theory and Applications. J Am Stat Assoc 1988;78:47-55.</p> <p>[22] McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) Frontiers in Health Policy Research. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.</p> <p>[23] Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. N Engl J Med. 2003;349:2117-2127.</p> <p>[20] Needleman, J., Buerhaus, P.I., Mattke, S., Stewart, M., and Zelevinsky, M. (2003). Health Services Research 38.6, Part I; 1487-1508.</p> <p>[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. Medical Care, 47(2): 226-233.</p>
<p>26</p> <p>(2c)</p>	<p>Validity Testing</p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the coronary artery bypass grafting surgery. We excluded small patient subgroups with much higher baseline risks, including those with procedure codes indicating that other operations were simultaneously performed (e.g., coronary artery bypass and valve surgery) or were performed for emergent indications.</p> <p>Analytic Method: We determined the value of our survival predictor (mortality) measure by establishing whether it explained hospital-level variation in risk-adjusted mortality rates and by assessing to what degree it was able to predict future hospital performance. We first estimated the proportion of variation in hospital-level mortality (2000-01) explained by the survival predictor measure using random effects logistic regression models. For these analyses, we estimated the proportional change in the hospital-level variance in mortality rates, which was determined from the standard deviation of the random effect, after adding each measure to the model [14,22].</p> <p>We next compared the ability of the survival predictor measure to the individual measures, mortality rates and hospital volume. We should note that these analyses focus on explaining systematic, or non-random,</p>

	<p>variation, since measurement error (random error) is accounted for and subtracted from the total variation in all analyses [22,24].</p> <p>We next determined the extent to which the composite measure predicts future risk-adjusted mortality. For this analysis, hospitals were ranked based on each measure from the earlier time period (data from years 2000-01) and divided into four equal size groups (quartiles at the patient level). The subsequent risk-adjusted mortality rates for each quartile of performance were then calculated (data from years 2002-03). We present the subsequent mortality rates across quartiles of the CABG survival predictor measure to graphically demonstrate its usefulness in discriminating among hospitals for the entire spectrum of performance. To compare the predictive ability of the composite measures and individual measures, we also present the subsequent mortality rates in the “worst” compared to the “best” quartile in the White Paper ([3], p. 22) "Quartiles of Performance Measures (2000-2001. This table reflects how well the unadjusted survival predictor created on 2000-2001 data compares to risk-adjusted mortality in 2002-2003 data.</p> <p>Note: The risk-adjusted mortality rate for CABG was constructed using standard methods. We determined the ratio of actual deaths or complications to the number of expected deaths (the O/E ratio). The number of expected deaths was the sum over all patients of the predicted probability of death or complications derived from a logistic regression model estimated on all patients undergoing CABG surgery. The dependent variable in the logistic model was death or complications and the independent variables were patient covariates. The patient characteristics included age, gender, race, admission acuity, and co-existing diseases using the Elixhauser method. A zip code level measure of socio-economic status was derived from 2000 census data.</p> <p>Testing Results: While some measures are good at discriminating top performers or bottom performers, this measure is good at prediction across entire spectrum of performance. [See White paper [3]: Figures p. 21-22) for a graphical demonstration of the usefulness of the survival predictor in discriminating among hospitals across the entire spectrum of performance.] To compare the predictive ability of the reliability adjusted survival predictor versus the individual components (volume and observed mortality) we also present the subsequent mortality rates in the "worst" compared to the "best" quartile.</p> <p>[22]. McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) Frontiers in Health Policy Research. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.</p> <p>[14] Birkmeyer JD, Dimick JB, Staiger DO. Operative mortality and procedure volume as predictors of subsequent hospital performance. Ann Surg 2006;243:411-417.</p> <p>[24] Zaslavsky AM, Cleary PD. Dimensions of plan performance for sick and healthy members on the Consumer Assessments of Health Plans Study 2.0 survey. Med Care 2002;40:951-964.</p>
<p>27 (2d)</p>	<p>Measure Exclusions <i>Provide evidence to justify exclusion(s) and analysis of impact on measure results during testing.</i></p> <p>Summary of Evidence supporting exclusion(s): The developers defined the denominator to minimize potential for case mix differences between hospitals, they excluded small patient sub-groups with much higher baseline risks, including those with procedure codes indicating that other operations were simultaneously performed. This essentially left a relatively homogenous population undergoing elective, non-emergency isolated surgeries. Hospitals that perform only emergent cases should not be compared to those performing elective surgeries. In addition, the denominator is based on isolated CABGs--excluding more complex surgeries such as: CABG and Valve combined.</p> <p>Only those hospitals with elective cases will have a survival predictor, since the primary goal of the measure is to provide information for selection of a specific hospital for the CABG procedure.</p> <p>Citations for Evidence:</p>

	<p>Data/sample:</p> <p>Analytic Method:</p> <p>Testing Results:</p>
28 (2e)	<p>Risk Adjustment Testing <i>Summarize the testing used to determine the need (or no need) for risk adjustment and the statistical performance of the risk adjustment method.</i></p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the coronary artery bypass grafting surgery. We excluded small patient subgroups with much higher baseline risks, including those with procedure codes indicating that other operations were simultaneously performed (e.g., coronary artery bypass and valve surgery) or were performed for emergent indications.</p> <p>Analytic Method: Sensitivity analysis. We performed a sensitivity analysis to determine whether risk-adjustment of the mortality input was important in improving the predictive ability of the survival predictor measure. Risk-adjustment was performed using logistic regression to estimate expected mortality rates for each hospital based on patient age, gender, race, urgency of operation, median income, and coexisting diseases. Coexisting diseases were determined from secondary diagnostic codes using the methods of Elixhauser (16). The observed mortality rate at each hospital was then divided by the expected mortality rate to yield the ratio of observed/expected deaths (O/E ratio). The O/E ratio was multiplied by the average mortality rate for each operation to yield a risk-adjusted mortality rate. To determine the value of risk-adjustment in the context of selective referral, we compared the ability of risk-adjusted and unadjusted composite measures to predict subsequent performance.</p> <p>Testing Results: In sensitivity analysis, composite measures based on an unadjusted mortality input and a risk-adjusted mortality input had a correlation of (.95) and thus were equally good a predicting future performance (See pages 21-22 in the White Paper [3]).</p> <p>► If outcome or resource use measure not risk adjusted, provide rationale: Because risk-adjusted mortality is not available publicly except for limited locations, the capacity to use unadjusted mortality is very desirable, especially since it was shown to provide (under this methodology) an equal result. This measure will allow measurement to occur across the United States, providing information to national companies, health plans and consumers.</p>
29 (2g)	<p>Testing comparability of results when more than 1 data method is specified (<i>e.g., administrative claims or chart abstraction</i>)</p> <p>Data/sample: not applicable</p> <p>Analytic Method:</p> <p>Results:</p>
30 (2f)	<p>Provide Measure Results from Testing or Current Use Results from testing</p> <p>Data/sample: same as described above, results for survival predictor in White Paper [3]available on Website and Validation results for composite in [16] Staiger, Dimick et al., Medical Care 2009</p> <p>Methods to identify statistically significant and practically/meaningfully differences in performance: Bayesian Hierarchical methods using new shrinkage estimator</p>

	<p>Empirical Bayesian methods to determine weights Correlations Calculated the amount of variation predicted by survival predictor as a percentage of all hospital-level variation (adjusted for sampling variation)--analogous to a R-squared from a regression that summarizes the ability of the predictor to explain the hospital level variation in mortality for CABG surgery. Predictor was tested against the "gold standard" --risk adjusted mortality</p> <p>Results: See White Paper [3]</p>
<p>31 (2h)</p>	<p>Identification of Disparities ▶ If measure is stratified by factors related to disparities (i.e. race/ethnicity, primary language, gender, SES, health literacy), provide stratified results: ▶ If disparities have been reported/identified, but measure is not specified to detect disparities, provide rationale: .</p>
<p>USABILITY</p>	
<p>32 (3)</p>	<p>Current Use Testing completed If in use, how widely used Nationally ▶ If "other," please describe: Survival Predictor for Pancreatectomy and Esophagectomy in use--see URL.</p> <p><input checked="" type="checkbox"/> Used in a public reporting initiative, name of initiative: Leapfrog Hospital Survey Sample report attached <input type="checkbox"/> OR Web page URL: https://www.leapfroggroup.org/cp</p>
<p>33 (3a)</p>	<p>Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)</p> <p>Data/sample:</p> <p>Methods:</p> <p>Results: See following citations reflecting consumer use of mortality information:</p> <p>[10] Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.</p> <p>[11] Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. Quality Review Bulletin 17:242-53.</p> <p>[12] Hibbard, J.H. and Jewett, J. (1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? Medical Care Research and Review., Vol 53(1): 28-47.</p>
<p>34 (3b, 3c)</p>	<p>Relation to other NQF-endorsed™ measures ▶ Is this measure similar or related to measure(s) already endorsed by NQF (on the same topic or the same target population)? <i>Measures can be found at www.qualityforum.org under Core Documents.</i> <i>Check all that apply</i></p> <p><input type="checkbox"/> Have not looked at other NQF measures <input checked="" type="checkbox"/> Other measure(s) on same topic <input checked="" type="checkbox"/> Other measure(s) for same target population <input type="checkbox"/> No similar or related measures</p> <p>Name and number of similar or related NQF-endorsed™ measure(s): STS/CMS CABG mortality volume; CA CABG Mortality</p> <p>Are the measure specifications harmonized with existing NQF-endorsed™ measures? Partially harmonized ▶ If not fully harmonized, provide rationale: This new measure requires combination of volume and mortality--no other measure uses this combination.</p> <p>Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: This measure provides the ability to produce reliable mortality results for low volume</p>

	hospitals, other measures do not have this capacity. In addition, the access to data nationally for other CABG mortality measures does not exist.
FEASIBILITY	
35 (4a)	<p>How are the required data elements generated? <i>Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data elements are generated concurrent with and as a byproduct of care processes during care delivery (e.g., blood pressure or other assessment recorded by personnel conducting the assessment)</p> <p><input type="checkbox"/> Data elements are generated from a patient survey (e.g., CAHPS)</p> <p><input checked="" type="checkbox"/> Data elements are generated through coding performed by someone other than the person who obtained the original information (e.g., DRG or ICD-9 coding on claims)</p> <p><input checked="" type="checkbox"/> Other, Please describe: Data are currently submitted to Leapfrog via a secure online survey-</p>
36 (4b)	<p>Electronic Sources All data elements</p> <p>► If all data elements are not in electronic sources, specify the near-term path to electronic collection by most providers:</p> <p>► Specify the data elements for the electronic health record: volume of CABG procedure, observed death during inpatient stay, related to CABG procedure</p>
37 (4c)	<p>Do the specified exclusions require additional data sources beyond what is required for the other specifications? No</p> <p>► If yes, provide justification:</p>
38 (4d)	<p>Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure: It is unlikely that this procedure, or inpatient death will be inaccurately coded or not coded given the high cost of procedure and the accompanying death.</p> <p>Describe how could these potential problems be audited: If problems were identified, a chart review of cases could be performed.</p> <p>Did you audit for these potential problems during testing? No If yes, provide results:</p>
39 (4e)	<p>Testing feasibility Describe what have you learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: Initial results only available for Esophagectomy, Pancreatectomy. CABG will be released in 2009</p>
CONTACT INFORMATION	
40	<p>Web Page URL for Measure Information <i>Describe where users (implementers) should go for more details on specifications of measures, or assistance in implementing the measure.</i></p> <p>Web page URL: https://leapfrog.medstat.com for access to Survival Predictor White Paper</p>
41	<p>Measure Steward Point of Contact</p> <p>First Name: MI: Last Name: Credentials (MD, MPH, etc.):</p> <p>Organization: The Leapfrog Group % The Academy</p> <p>Street Address: 1150 17th St., NW, Suite 600 City: Washington State: DC ZIP: 20036</p> <p>Email: Telephone: ext:</p>
42	<p>Measure Developer Point of Contact If different from Measure Steward</p> <p>First Name: Justin MI: B Last Name: Dimick Credentials (MD, MPH, etc.): MD, MPH</p> <p>Organization: Department of Surgery, University of Michigan, M-SCORE offices, Suite 201 and 202</p> <p>Street Address: 211 N. Fourth Avenue City: Ann Arbor State: MI ZIP: 48104</p> <p>Email: jdimick@umich.edu Telephone: ext:</p>
ADDITIONAL INFORMATION	
43	<p>Workgroup/Expert Panel involved in measure development Workgroup/panel used</p> <p>► If workgroup used, describe the members' role in measure development: Research team led by Justin</p>

	<p>Dimick, MD, MPH; ► Provide a list of workgroup/panel members' names and organizations:</p> <p>Douglas Staiger Ph.D., Department of Economics and the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth College, Hanover, New Hampshire</p> <p>John D. Birkmeyer, MD Michigan Surgical Collaborative for Outcomes Research and Evaluation Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Onur Baser, Ph.D. Michigan Surgical Collaborative for Outcomes Research and Evaluation Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Research supported by the National Institute on Aging</p>
44	<p>Measure Developer/Steward Updates and Ongoing Maintenance Year the measure was first released: 2008 Month and Year of most recent revision: August 2008 What is the frequency for review/update of this measure? Annual When is the next scheduled review/update for this measure? New coefficients for August 2009</p>
45	<p>Copyright statement/disclaimers: none</p>
46	<p>Additional Information: All measure information is available at https://leapfrog.medstat.com Please contact measure developer prior to use to assure all necessary items have been accessed.</p>
47	<p>I have checked that the submission is complete and any blank fields indicate that no information is provided. <input checked="" type="checkbox"/></p>
48	<p>Date of Submission (MM/DD/YY): Revised submission dated 3/18/09</p>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

The measure information you submit will be shared with NQF’s Steering Committees and Technical Advisory Panels to evaluate measures against the NQF criteria of importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Four conditions (as indicated below) must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, *importance to measure and report*, in order to be evaluated against the remaining criteria. References to the specific measure evaluation criteria are provided in parentheses following the item numbers. Please refer to the *Measure Evaluation Criteria* for more information at www.qualityforum.org under Core Documents. Additional guidance is being developed and when available will be posted on the NQF website.

Use the tab or arrow (↓→) keys to move the cursor to the next field (or back ←↑). There are three types of response fields:

- drop-down menus - select one response;
- check boxes - check as many as apply; and
- text fields - you can copy and paste text into these fields or enter text; these fields are not limited in size, but in most cases, we ask that you summarize the requested information.

Please note that URL hyperlinks do not work in the form; you will need to type them into your web browser.

Be sure to answer all questions. Fields that are left blank will be interpreted as no or none. **Information must be provided in this form.** Attachments are not allowed except to provide additional detail or source documents for information that is summarized in this form. If you have important information that is not addressed by the questions, they can be entered into item #46 near the end of the form.

For questions about this form, please contact the NQF Project Director listed in the corresponding call for measures.

CONDITIONS FOR CONSIDERATION BY NQF	
	Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards.
A (A)	Public domain or Measure Steward Agreement signed: Public domain - Agreement not required (If no, do not submit) <i>Template for the Measure Steward Agreement is available at www.qualityforum.org under Core Documents.</i>
B (B)	Measure steward/maintenance: Is there an identified responsible entity and process to maintain and update the measure on a schedule commensurate with clinical innovation, but at least every 3 years? Yes, information provided in contact section (If no, do not submit)
C (C)	Intended use: Does the intended use of the measure include BOTH public reporting AND quality improvement? Yes (If no, do not submit)
D (D)	Fully developed and tested: Is the measure fully developed AND tested? Yes, fully developed and tested (If not tested and no plans for testing within 24 months, do not submit)

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

	(for NQF staff use) NQF Review #: HOE-020-08	NQF Project: Hospital Outcomes and Efficiency
MEASURE SPECIFICATIONS & DESCRIPTIVE INFORMATION		
1	Information current as of (date- MM/DD/YY):	
2	Title of Measure: Survival Predictor for Percutaneous Coronary Interventions (PCI)	
3	Brief description of measure ¹ : A reliability adjusted measure of PCI performance that optimally combines two important domains: PCI hospital volume and PCI operative mortality, to provide predictions on PCI survival rates for hospitals. This measure is calculated based on data from administrative claims information.	
4 (2a)	<p>Numerator Statement: Note: Because of the type of modeling done for this Survival Predictor--the information is not readily split into Numerator/ Denominator statements. Thus, we describe the two domains and their coding and data needs in this section.</p> <p>The formula for calculating the survival predictor has two components, one is a volume predicted mortality rate, and the second is an observed mortality rate. The volume predicted mortality rate reflects the hospitals experience performing PCI surgeries (thus, it includes all PCI surgeries) and uses mortality for all hospitals at that specific volume to create the volume predicted mortality. The input data from the hospitals for this domain is a volume count of all PCIs performed in the hospital.</p> <p>The second domain is the observed mortality, for this domain the population is the group of PCI cases, the data needed for this domain is the number of observed deaths occurring for PCI cases, within the inpatient setting.</p> <p>Note: All data is available in administrative claims information. In the case of Leapfrog's implementation hospitals are asked to submit aggregated information from their claims data. No personal health information is submitted to Leapfrog. Other users of the measure may have direct access to administrative data.</p> <p>Time Window: T</p> <p>Numerator Details (Definitions, codes with description): For the volume predicted mortality, hospitals count the number of PCI cases using the following codes:</p> <p>ICD-9-CM Procedure</p> <ul style="list-style-type: none"> ■ 00.66 Percutaneous transluminal coronary angioplasty (PTCA) or coronary atherectomy ■ 36.01 Single vessel percutaneous transluminal coronary angioplasty without mention of thrombolytics (code discontinued 10/1/2005) ■ 36.02 Single vessel percutaneous transluminal coronary angioplasty with mention of thrombolytics (code discontinued 10/1/2005) ■ 36.05 36.05 Multiple vessel PTCA at the same session with or without mention of thrombolytics (code discontinued 10/1/2005) ■ 36.06 Insertion of non-drug eluting coronary stents ■ 36.07 insertion of drug eluting coronary stents 	

¹ Example of measure description: Percentage of adult patients with diabetes aged 18-75 years receiving one or more A1c test(s) per year.
NQF Measure Submission Form, V3.1

	<p>See calculation worksheet for details on how volume-predicted mortality is used in the model.</p> <p>For the observed mortality domain, the hospital submits the observed deaths for PCI cases using the following codes (NQF endorsed)</p> <ul style="list-style-type: none"> ■ 00.66 Percutaneous transluminal coronary angioplasty (PTCA) or coronary atherectomy ■ 36.01 Single vessel percutaneous transluminal coronary angioplasty without mention of thrombolytics (code discontinued 10/1/2005) ■ 36.02 Single vessel percutaneous transluminal coronary angioplasty with mention of thrombolytics (code discontinued 10/1/2005) ■ 36.05 36.05 Multiple vessel PTCA at the same session with or without mention of thrombolytics (code discontinued 10/1/2005) ■ 36.06 Insertion of non-drug eluting coronary stents ■ 36.07 insertion of drug eluting coronary stents <p>See Calculation Worksheet for examples of how the two domains are used to create the Survival Predictor.</p>
5 (2a)	<p>Denominator Statement:</p> <p>Time Window:</p> <p>Denominator Details (Definitions, codes with description):</p>
6 (2a, 2d)	<p>Denominator Exclusions: No exclusions)</p> <p>Denominator Exclusion Details (Definitions, codes with description): (</p>
7 (2a, 2h)	<p>Stratification Do the measure specifications require the results to be stratified? No ► If "other" describe:</p> <p>Identification of stratification variable(s):</p> <p>Stratification Details (Definitions, codes with description):</p>
8 (2a, 2e)	<p>Risk Adjustment Does the measure require risk adjustment to account for differences in patient severity before the onset of care? No ► If yes, (select one) ► Is there a separate proprietary owner of the risk model? No</p> <p>Identify Risk Adjustment Variables: See section 28 for rationale and support for not risk adjusting this measure. Measure was tested against risk adjusted mortality--details on that provided in Section 26.</p> <p>Detailed risk model: attached <input type="checkbox"/> OR Web page URL:</p>
9 (2a)	<p>Type of Score: Rate/proportion Calculation Algorithm: attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p>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 = Score within a defined interval ► If "Other", please describe:</p>
10 (2a, 4a, 4b)	<p>Identify the required data elements(e.g., primary diagnosis, lab values, vital signs): procedure codes Data dictionary/code table attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p>Data Quality (2a) Check all that apply</p> <ul style="list-style-type: none"> <input checked="" type="checkbox"/> Data are captured from an authoritative/accurate source (e.g., lab values from laboratory personnel) <input checked="" type="checkbox"/> Data are coded using recognized data standards <input checked="" type="checkbox"/> Method of capturing data electronically fits the workflow of the authoritative source <input type="checkbox"/> Data are available in EHRs <input checked="" type="checkbox"/> Data are auditable

<p>11 (2a, 4b)</p>	<p>Data Source and Data Collection Methods <i>Identifies the data source(s) necessary to implement the measure specifications. Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Electronic Health/Medical Record</td> <td><input type="checkbox"/> Paper Medical Record</td> </tr> <tr> <td><input type="checkbox"/> Electronic Clinical Database, Name:</td> <td><input type="checkbox"/> Standardized clinical instrument, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Clinical Registry, Name:</td> <td><input type="checkbox"/> Standardized patient survey, Name:</td> </tr> <tr> <td><input checked="" type="checkbox"/> Electronic Claims</td> <td><input type="checkbox"/> Standardized clinician survey, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Pharmacy data</td> <td><input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.</td> </tr> <tr> <td><input type="checkbox"/> Electronic Lab data</td> <td></td> </tr> <tr> <td><input type="checkbox"/> Electronic source - other, Describe:</td> <td></td> </tr> </table> <p style="text-align: right;">Instrument/survey attached <input checked="" type="checkbox"/> OR Web page URL:</p>	<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record	<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:	<input type="checkbox"/> Electronic Clinical Registry, Name:	<input type="checkbox"/> Standardized patient survey, Name:	<input checked="" type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:	<input type="checkbox"/> Electronic Pharmacy data	<input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.	<input type="checkbox"/> Electronic Lab data		<input type="checkbox"/> Electronic source - other, Describe:					
<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record																		
<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:																		
<input type="checkbox"/> Electronic Clinical Registry, Name:	<input type="checkbox"/> Standardized patient survey, Name:																		
<input checked="" type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:																		
<input type="checkbox"/> Electronic Pharmacy data	<input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.																		
<input type="checkbox"/> Electronic Lab data																			
<input type="checkbox"/> Electronic source - other, Describe:																			
<p>12 (2a)</p>	<p>Sampling <i>If measure is based on a sample, provide instructions and guidance on sample size.</i> Minimum sample size: Instructions:</p>																		
<p>13 (2a)</p>	<p>Type of Measure: Outcome ▶ If "Other", please describe:</p> <p>▶ If part of a composite or paired with another measure, please identify composite or paired measure While the measure uses two types of information components (domains), the results are not a composite as is defined by NQF, but rather a reliability adjusted measure of survival.</p> <p>Volume is used to create a volume predicted mortality for the hospital--this component of the measure is used to create greater reliability for low-volume hospitals. In the modeling for this measure, the volume predicted mortality and the observed mortality are weighted. In the model, lower volume hospitals have a higher weight on the volume predicted mortality versus the observed mortality. The opposite is true for high volume hospitals, which have a higher weight on the observed mortality. This methodology results in a reliability adjusted survival predictor.</p>																		
<p>14 (2a)</p>	<p>Unit of Measurement/Analysis <i>(Who or what is being measured) Check all that apply.</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Can be measured at all levels</td> <td><input type="checkbox"/> Integrated delivery system</td> </tr> <tr> <td><input type="checkbox"/> Individual clinician (e.g., physician, nurse)</td> <td><input checked="" type="checkbox"/> Health plan</td> </tr> <tr> <td><input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)</td> <td><input checked="" type="checkbox"/> Community/Population</td> </tr> <tr> <td><input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)</td> <td><input type="checkbox"/> Other <i>(Please describe):</i></td> </tr> </table>	<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system	<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input checked="" type="checkbox"/> Health plan	<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input checked="" type="checkbox"/> Community/Population	<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other <i>(Please describe):</i>										
<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system																		
<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input checked="" type="checkbox"/> Health plan																		
<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input checked="" type="checkbox"/> Community/Population																		
<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other <i>(Please describe):</i>																		
<p>15 (2a)</p>	<p>Applicable Care Settings <i>Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Can be used in all healthcare settings</td> <td><input type="checkbox"/> Hospice</td> </tr> <tr> <td><input type="checkbox"/> Ambulatory Care (office/clinic)</td> <td><input checked="" type="checkbox"/> Hospital</td> </tr> <tr> <td><input type="checkbox"/> Behavioral Healthcare</td> <td><input type="checkbox"/> Long term acute care hospital</td> </tr> <tr> <td><input type="checkbox"/> Community Healthcare</td> <td><input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)</td> </tr> <tr> <td><input type="checkbox"/> Dialysis Facility</td> <td><input type="checkbox"/> Prescription Drug Plan</td> </tr> <tr> <td><input type="checkbox"/> Emergency Department</td> <td><input type="checkbox"/> Rehabilitation Facility</td> </tr> <tr> <td><input type="checkbox"/> EMS emergency medical services</td> <td><input type="checkbox"/> Substance Use Treatment Program/Center</td> </tr> <tr> <td><input type="checkbox"/> Health Plan</td> <td><input type="checkbox"/> Other <i>(Please describe):</i></td> </tr> <tr> <td><input type="checkbox"/> Home Health</td> <td></td> </tr> </table>	<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice	<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital	<input type="checkbox"/> Behavioral Healthcare	<input type="checkbox"/> Long term acute care hospital	<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)	<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan	<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility	<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center	<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other <i>(Please describe):</i>	<input type="checkbox"/> Home Health	
<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice																		
<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital																		
<input type="checkbox"/> Behavioral Healthcare	<input type="checkbox"/> Long term acute care hospital																		
<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)																		
<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan																		
<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility																		
<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center																		
<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other <i>(Please describe):</i>																		
<input type="checkbox"/> Home Health																			
<p>IMPORTANCE TO MEASURE AND REPORT</p>																			
<p>Note: This is a threshold criterion. If a measure is not judged to be sufficiently important to measure and report, it will not be evaluated against the remaining criteria.</p>																			
<p>16 (1a)</p>	<p>Is measure related to a National Priority Partners priority area? Safety reliability (for NQF staff use) Does measure address a <u>specific</u> NPP goal? (www.qualityforum.org/about/NPP/):</p>																		
<p>17</p>	<p>Does the measure address a high impact aspect of healthcare patient/societal consequences of poor quality</p>																		

(1a)

Summary of Evidence: This measure addresses mortality in a high risk procedure (PCI) and is an outcome measure which is of interest to both consumers and purchasers.

Rathore et.al., (2004) [1a] examined mortality rates for PCI and found that crude in-hospital rates were 2.56% in low volume hospitals, 1.83% in medium-volume hospitals, and 1.64% in high volume hospitals, and 1.36% in very high volume hospitals ($p < 0.001$ for trend). While the actual rate of mortality is low compared to some surgeries, the volume of cases is very high. For example, in one year in New York Hospitals there were just under 70,000 PCIs performed. [1b]. Expenditures for AMI which is the principal diagnosis in many PCIs was in the top 5 diagnoses for Medicare patients in 2006. Expenditures for AMI were over \$19 billion in 2006 for Medicare patients [2]. AMI is number 4 in the top 20 most expensive conditions treated in US hospitals. [2]

In addition to addressing high volume procedure risk, this measure improves upon the technology of surgical procedure mortality measurement. It overcomes three problems with existing PCI mortality measures: 1) Mortality rates are often too "noisy" to reflect hospital quality with surgery (particularly among lower volume hospitals), 2) volume alone is a weak proxy for most procedures, and 3) when both volume and mortality are reported as separate indicators it is difficult to understand which measure is more important. [1] Given the large number of PCI procedures performed annually in the United States, and that this measure specifically addresses hospitals which perform elective procedures, consumers and purchasers would benefit from information that is more reliable in the prediction of future mortality for both selection and selective referral. In addition, this measure can be applied to the nation, states, or regions.

Birkmeyer and Dimick (2009)[4] show that differences in mortality can be predicted using a reliability adjusted mortality rate (a weighted combination of volume and mortality) which is particularly relevant for selective-referral or public reporting contexts. They reduce the effects of random chance (statistical noise) and as a result with CABG, for example, more than half of the observed variation can be attributed to statistical noise. When they sorted hospitals simply on actual (risk-adjusted) mortality, rates varied from 1.4% to 11.0% across hospital quintiles (Figure 1 in White Paper [3]). After they adjusted for reliability, however, the mortality rates varied considerably less, from 3.3% to 6.3%. Given Rathore et al., findings of a significant trend difference at various levels of volume, it is clear that there is an opportunity and room for improvement.

Citations² for Evidence:

[1a] Rathore, A.J., Epstein, A.J., Rathore, S.S., Volpp, K.G., and Krunholz, H.M. (2004). Hospital percutaneous coronary intervention volume and patient mortality. 1988 to 2000: does the evidence support current procedure volume minimums? *J Am Coll Surg.*; 43:(10): 1755-62.

[1b] Moscucci, M., Eagle, K.A., Share, D., Smith, D., DeFranco, A.C., O'Donnell, M., Kline-Rogers, E., Jani, S.M., and Brown, D.L. (2005). Public Reporting and Case Selection for Percutaneous Coronary Interventions: An Analysis from Two Large Multicenter Percutaneous Coronary Intervention Databases. *J. Am Coll Card.*, 45(11):1759-1765.

[2] The National Hospital Bill: The Most Expensive Conditions by Payer, 2006. Statistical Brief #59. File accessed on March 16, 2009, at: <http://www.hcup-us.ahrq.gov/reports/statbriefs/sb59.jsp> Produced by AHRQ, Center for Delivery, Organization, and Markets, Healthcare Cost and Utilization Project, Nationwide Inpatient Sample, 2006.

[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: <http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf>

[4] Birkmeyer, J.D., and Dimick, J.B. (2009) Understanding and reducing variation in surgical mortality. *Annu. Rev. Med.* 2009. 60:405-15.

² Citations can include, but are not limited to journal articles, reports, web pages (URLs).
NQF Measure Submission Form, V3.1

<p>18 (1b)</p>	<p>Opportunity for Improvement <i>Provide evidence that demonstrates considerable variation, or overall poor performance, across providers.</i></p> <p>Summary of Evidence: In 2002, a systematic review of the literature on the volume-outcome relationship found that there was a significant relationship between hospital volume and outcomes for PCI surgery. While this relationship was not as robust in PCI procedures as it was in some other surgical procedures (esophagectomy, pancreatectomy), it was present. [5]</p> <p>As indicated in Section 17, the volume relationship with mortality indicates variation as shown by the linear trend data, [1a] indicates that variation exists across providers and that it is clearly related to volume of procedures.</p> <p>Given the findings related to volume of procedures, Silber et al., [7] explored the relative contribution of complication rates and failure to rescue rates to mortality and found that complication rates were more likely influenced by patient factors while failure to rescue rates of those with complications was more related to hospital factors. Thus, it may be that higher volume hospitals are better at rescuing patients with complications. Silbers finding, in conjunction with the volume information, suggests lower volume hospitals with worse mortality rates could in fact address this through better care following the procedure, thereby reducing their overall rate. Unfortunately, most low volume hospitals in the United States do not have information on their PCI mortality rate compared to other hospitals. When they are given this information, there is a good chance for improvement. Birkmeyer and Dimick [4] report that in northern New England, mortality associated with CABG fell by >25% when hospitals and surgeons were given feedback on their mortality data.</p> <p>Note: Birkmeyer and Dimmick [4] indicate it is also likely that some lower volume hospitals would also have lower mortality rates.</p> <p>Citations for Evidence:</p> <p>[1a] Rathore, A.J., Epstein, A.J., Rathore, S.S., Volpp, K.G., and Krunholz, H.M. (2004). Hospital percutaneous coronary intervention volume and patient mortality. 1988 to 2000: does the evidence support current procedure volume minimums? J Am Coll Surg.; 43:(10): 1755-62.</p> <p>[5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. Annals of Internal Medicine, Sept 1;137(6):511-20</p> <p>[6] Epstein, A.J., Rathore, S.S., Krumholz, H.M., and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. BMC Health Services Research, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>[7] Silber, J.H., Rosenbaum, P.R., Trudeau, M.E., et al. 2005. Changes in prognosis after the first postoperative complication. Medical Care, 43:122-31.</p>
<p>19 (1b)</p>	<p>Disparities <i>Provide evidence that demonstrates disparity in care/outcomes related to the measure focus among populations.</i></p> <p>Summary of Evidence: It is more likely that minorities will be treated at a low volume facility, and as a result are likely to be impacted by higher mortality rates. In an analysis of the National Inpatient Sample, Epstein, Rathore and Krumholz (2005)[6] found that a greater proportion of patients treated in low volume hospitals for cardiovascular conditions were non-white, while a lower proportion of non-white patients presented as "elective" admissions or patients received in transfer as compared to patients in high volume hospitals.</p> <p>Citations for evidence:</p> <p>[6, p. 3-5]</p>

<p>20 (1c)</p>	<p>If measuring an Outcome Describe relevance to the national health goal/priority, condition, population, and/or care being addressed: A PCI procedure is a frequently performed procedure with some risk of mortality. It is performed on patients with AMI; those patients with this diagnosis constitute the number 4 condition being treated in the US for Medicare patients. Given the risk associated with the procedure, the national expenditure, and some limited information on potential overuse, the Survival Predictor provides the first national information on patient outcomes. Other entities with clinical information are not publicly reporting mortality rates of PCI procedures by hospital provider. Only New York state reports on the number of PCIs performed and the associated mortality. Prominent private registries are not producing public information, e.g., American College of Cardiology, or the Society for Thoracic Surgery. This measure is designed to give feedback to hospitals across the country as well as to provide information for decision-making by consumers and purchasers.</p> <p>Mortality in US hospitals varies for PCI--there are documented differences between high and low performing hospitals [4]. Higher volumes are associated with better outcomes including lower mortality. In addition to being a high risk surgery, this surgery is one of the high cost procedures by total bill in this country. In 2006, more than \$19 billion in Medicare payments were spent on AMI of which PCI made up a significant proportion, this is 4.3% of the national annual bill for hospitals; it also accounts for 430,000 hospital stays. It is number 4 in the top 20 most expensive conditions for Medicare treated in US hospitals. [2]</p> <p>This measure is highly relevant to both consumers and purchasers, given its frequency. National purchasers are interested in comparative information on hospitals nationwide. Pauly (1996) in a study of purchaser interests in hospital performance reporting found that mortality ratings were more important to purchasers than were morbidity or complications. [9] Health plans are interested in contracting with centers of excellence, which can be identified through the results of survival predictor in combination with other information on cost and quality. Consumers have shown their interest in cardiac procedure mortality by requesting reports from the state of Pennsylvania [10]; an earlier study by IOM (Lohr, Donaldson and Walker 1991) found that consumers were interested in hospital mortality rates, but did not perceive this information to be available.[11] Hibbard and Jewett found that consumers were more interested in "undesirable events" (such as mortality, complications, infections) than in "desirable events." [12]</p> <p>[9] Pauly, M.V., Brailer, D.J.Kroch, E., and Even-Shoshan, O. Measuring Hospital Outcomes from a Buyer's Perspective. American Journal of Medical Quality, 11(8): Fall 1996.</p> <p>[10] Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.</p> <p>[11]Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. Quality Review Bulletin 17:242-53.</p> <p>[12]Hibbard, J.H. and Jewett, J.(1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? Medical Care Research and Review., Vol 53(1): 28-47.</p> <p>If not measuring an outcome, provide evidence supporting this measure topic and grade the strength of the evidence <i>Summarize the evidence (including citations to source) supporting the focus of the measure as follows:</i></p> <ul style="list-style-type: none"> • <u>Intermediate outcome</u> - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit. • <u>Process</u> - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s). • <u>Structure</u> - evidence that the measured structure supports the consistent delivery of effective

	<p>processes or access that lead to improved health/avoidance of harm or cost/benefit.</p> <ul style="list-style-type: none"> • <u>Patient experience</u> - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public. • <u>Access</u> - evidence that an association exists between access to a health service and the outcomes of, or experience with, care. • <u>Efficiency</u>- demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality. <p>Type of Evidence <i>Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Evidence-based guideline</td> <td><input checked="" type="checkbox"/> Quantitative research studies</td> </tr> <tr> <td><input checked="" type="checkbox"/> Meta-analysis</td> <td><input type="checkbox"/> Qualitative research studies</td> </tr> <tr> <td><input checked="" type="checkbox"/> Systematic synthesis of research</td> <td><input type="checkbox"/> Other (<i>Please describe</i>):</td> </tr> </table> <p>Overall Grade for Strength of the Evidence³ (<i>Use the USPSTF system, or if different, also describe how it relates to the USPSTF system</i>): Moderate</p> <p>Summary of Evidence (<i>provide guideline information below</i>): Over 100 articles published related to volume and outcome relationship, with some inconsistency in results. Systematic review of the literature conducted in 2002. No review since that time.</p> <p>Citations for Evidence:</p> <p>[5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. <i>Annals of Internal Medicine</i>, Sept 1;137(6):511-20.</p> <p>[14] Birkmeyer, J.D., Dimick, J.B., Staiger, D.O. (2006) Operative mortality and procedure volume as predictors of subsequent hospital performance. <i>Ann Surg</i>. 243:411-417.</p> <p>[15] Dimick, JB, Welch HG, Birkmeyer JD. (2004) Surgical mortality as an indicator of hospital quality: The problem with small sample size. <i>JAMA</i>, 292:847-851.</p> <p>[4] Birkmeyer, JD., and Dimick, JB. (2009) Understanding and reducing variation in surgical mortality. <i>Annu. Rev. Med.</i> 60:405-15.</p> <p>[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. <i>Medical Care</i>, 47(2): 226-233.</p> <p>[6] Andrew J Epstein, Saif S Rathore, Harlan M Krumholz and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. <i>N Engl J Med</i>. 1979;301:1364-9.</p>	<input type="checkbox"/> Evidence-based guideline	<input checked="" type="checkbox"/> Quantitative research studies	<input checked="" type="checkbox"/> Meta-analysis	<input type="checkbox"/> Qualitative research studies	<input checked="" type="checkbox"/> Systematic synthesis of research	<input type="checkbox"/> Other (<i>Please describe</i>):
<input type="checkbox"/> Evidence-based guideline	<input checked="" type="checkbox"/> Quantitative research studies						
<input checked="" type="checkbox"/> Meta-analysis	<input type="checkbox"/> Qualitative research studies						
<input checked="" type="checkbox"/> Systematic synthesis of research	<input type="checkbox"/> Other (<i>Please describe</i>):						
<p>21 (1c)</p>	<p>Clinical Practice Guideline <i>Cite the guideline reference; quote the specific guideline recommendation related to the measure and the guideline author's assessment of the strength of the evidence; and summarize the rationale for using this guideline over others.</i></p> <p>Guideline Citation: Ryan TJ, Bauman, WB., Kennedy, JW., et al., Guidelines for percutaneous</p>						

³The strength of the body of evidence for the specific measure focus should be systematically assessed and rated, e.g., USPSTF grading system www.ahrq.gov/clinic/uspstmeth.htm: **A** - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. **B** - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. **C** - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. **D** - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. **I** - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

	<p>transluminal coronary angioplasty: a report of the American Heart Association /American College of Cardiology Task Force on Assessment of Diagnostic and Therapeutic Cardiovascular Procedures (Committee on Percutaneous Transluminal Coronary Angioplasty) <i>Circulation</i> (1993) 88, 2987-3007.</p> <p>Specific guideline recommendation: Guidelines recommend that physicians perform at least 75 procedures and hospitals perform at least 400 procedures annually.</p> <p>Guideline author's rating of strength of evidence (<i>If different from USPSTF, also describe it and how it relates to USPSTF</i>):</p> <p>Rationale for using this guideline over others: Not aware of other guidelines on this topic of volume for combined physician and hospital volume standards.</p>
<p>22 (1c)</p>	<p>Controversy/Contradictory Evidence Summarize any areas of controversy, contradictory evidence, or contradictory guidelines and provide citations.</p> <p>Summary: There are three areas of possible contention with this measure--</p> <p>1) The volume-outcome relationship has been questioned for some procedures [6, 17, 19, 25]</p> <p>Peterson et al., [19] questioned the volume outcomes relationship for CABG surgery, and found only modest associations for volume and outcome for CABG. Those with high volume had mortality rate of 2.5% while low volume hospitals rate was 3.2%. They suggest using past mortality rate to select hospitals. (The survival predictor uses both volume and mortality to predict survival in the next year.)</p> <p>Yet, more than 100 studies have demonstrated better results at high-volume hospitals with cardiovascular surgery, major cancer resections, and other high-risk procedures.[18, 20]</p> <p>There is all specific evidence of the variation in mortality across the different volume levels of PCIs. [1a]. They documented that there were differences between low volume hospitals and high volume hospitals in mortality--with high volume hospitals having less mortality.</p> <p>McGrath et al., 2000, concluded that Medicare patients treated by high-volume physicians and at high-volume centers(hospitals) that patients experience better outcomes. Their findings support the ACC guideline on physician and hospital volume for PCIs.</p> <p>2) That outcome measures must be risk-adjusted unless there is evidence to show it is not needed (NQF). The survival predictor measure predicts better than volume or mortality alone, and is as good a predictor as risk-adjusted mortality. When testing the unadjusted survival predictor against risk-adjusted mortality there was a (.96) correlation. [4] See Section 28 of this form for details.</p> <p>3) The weighting of input measures into composites. Existing approaches rely on overly simplistic approaches. Among these, assigning equal weight to all measures (i.e., the all or none approach) and relying on expert opinion are the most common. The survival predictor relies on empiric methods for weighting the input measures.</p> <p>Citations:</p> <p>[1a] Rathore, A.J., Epstein, A.J., Rathore, S.S., Volpp, K.G., and Krumholz, H.M. (2004). Hospital percutaneous coronary intervention volume and patient mortality. 1988 to 2000: does the evidence support current procedure volume minimums? <i>J Am Coll Surg.</i>; 43:(10): 1755-62.</p> <p>[6] Andrew J Epstein, Saif S Rathore, Harlan M Krumholz and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>[16]Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. <i>Medical Care</i>, 47(2): p. 232.</p>

	<p>[17] Edward L. Hannan, PhD; Chuntao Wu, PhD; Thomas J. Ryan, MD; Edward Bennett, MD; Alfred T. Culliford, MD; Jeffrey P. Gold, MD; Alan Hartman, MD; O. Wayne Isom, MD; Robert H. Jones, MD; Barbara McNeil, MD, PhD; Eric A. Rose, MD; Valavanur A. Subramanian, MD. Do Hospitals and Surgeons With Higher Coronary Artery Bypass Graft Surgery Volumes Still Have Lower Risk-Adjusted Mortality Rates? <i>Circulation</i>. 2003;108:795-801.</p> <p>[18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. <i>N Engl J Med</i>. 1979;301:1364-9.</p> <p>[19] Eric D. Peterson, MD, MPH; Laura P. Coombs, PhD; Elizabeth R. DeLong, PhD; Constance K. Haan, MD; T. Bruce Ferguson, MD. Procedural Volume as a Marker of Quality for CABG Surgery. <i>JAMA</i>. 2004;291:195-201.</p> <p>[20] Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. <i>JAMA</i>. 1998;280:1747-51.)</p> <p>[25] McGrath, PD., Wennberg, DE., Dickens, Jr., JD., Siewers, AE., Lucas, FL., Malenka, DJ, Kellett, Jr., MA., Ryan, Jr., TJ. (2000) Relation Between Operator and Hospital Volume and Outcomes Following Percutaneous Coronary Interventions in the Era of the Coronary Stent. <i>JAMA</i>, 284(24):3139-3144.</p>
<p>23 (1)</p>	<p>Briefly describe how this measure (as specified) will facilitate significant gains in healthcare quality related to the specific priority goals and quality problems identified above: This measure of predicted survival improves upon the reliability of mortality results for high risk surgical procedures, such as PCI. For the first time, this measure produces reliable mortality/survivability information on smaller volume hospitals, as well as high volume hospitals. Hospitals across the country will have information available through voluntary public reporting.</p>
<p>SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES</p>	
<p>Note: Testing and results should be summarized in this form. However, additional detail and reports may be submitted as supplemental information or provided as a web page URL. If a measure has not been tested, it is only potentially eligible for time-limited endorsement.</p>	
<p>24</p>	<p>Supplemental Testing Information: attached <input checked="" type="checkbox"/> OR Web page URL:</p>
<p>25 (2b)</p>	<p>Reliability Testing</p> <p>Data/sample: Data was a 100% sample from the Medicare Analysis Provider and Review (MEDPAR) files for 2000-2003, these files contain 100% of Medicare hospitalizations for years specified. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing a percutaneous coronary intervention. Codes used to select patients are indicated in Section 4 of this form.</p> <p>This dataset has been used by other researchers to look at the issue of volume and mortality for PCI. As indicated by McGrath and Wennberg et al., [25] the Medicare dataset allowed sufficient power to determine significant differences in adverse outcomes across varying levels of volume.</p> <p>Note: Needleman, Buerhaus, et al., (2003) concluded after applying operational tests on Medicare data for adverse outcomes and all-patient hospital data from 11 states, that Medicare data could be used to assess quality in hospitals.[20] Given the lack of a national all-patient database, MEDPAR data was used in development and testing of the models.</p>

Analytic Method:**Model Development**

We used an empirical Bayes approach to combine mortality rates with information on hospital volume at each hospital. In traditional empirical Bayes methods, a point estimate (e.g., mortality rate observed at a hospital) is adjusted for reliability by shrinking it towards the overall mean (e.g., overall mortality rate in the population) [21,22]. We modified this traditional approach by shrinking the observed mortality rate back toward the mortality rate expected given the volume at that hospital—we refer to this as the “volume-predicted mortality” (See attached White Paper [3] TECHNICAL APPENDIX for the mathematical details of this method). With this approach, the observed mortality rate is weighted according to how reliably it is estimated, with the remaining weight placed on the information regarding hospital volume. Because this method includes observed data to the extent that it is useful, and only relies on the proxy measure to the extent necessary, it ensures an optimal combination of these two quality domains. [3]

The two inputs to the survival predictor measure are mortality rates and procedure volume for each of the six included operations. Procedure-specific mortality rates were calculated for all hospitals over a 2-year period (2000-01) and this was used as the first input. Hospital volume was calculated as the number of Medicare cases performed during the same time period. For each operation, the relationship between hospital volume and risk-adjusted mortality was modeled using linear regression. (Details of the risk-adjustment strategy will be discussed below.) After testing the fit of several transformations, hospital volume was modeled as the natural log of the continuous volume variable, which is the same approach used in our previous work [23]. Using this regression model, we estimated the volume-predicted mortality, the second input to the survival predictor measure.

We then used the empirical Bayes approach to create an optimal combination of these two inputs. This survival predictor measure theoretically provides the best estimate of a hospital's true mortality rate, taking into account the both available inputs [21,22]. The combined survival predictor measure was calculated as follows: mortality prediction = (weight)*(observed mortality) + (1-weight)*(volume-predicted mortality). The weight placed on the point estimate of mortality is the reliability, or ratio of signal to signal plus noise, calculated as follows: weight = variation among hospitals/(variation among hospitals + variation within hospitals). The variation among hospitals was calculated as the variance in observed mortality rates for the hospitals included in the sample. The variation within hospitals was calculated as the standard error of the mortality rate at each hospital. With this method, more weight is placed on the observed mortality rate when a hospital has a high number of cases because it is estimated with more reliability; less weight is placed on the observed mortality rate when a hospital performs a low number of cases because of its lower reliability.

A calculation worksheet with examples is attached.

Testing Results: Hospital caseloads and the weights applied to each input to the survival predictor measure varied for each procedure studied (see Table 1 in white paper [3]). For percutaneous coronary intervention, a procedure with relatively high hospital caseloads, the weight applied to the volume input was .52. ([3]-Table 1). For hospitals with higher volumes more weight was placed on the observed mortality.

The survival predictor (mortality) measure explained a large proportion of non-random, hospital-level variation in risk-adjusted mortality rates (see Table 2, p. 19 in White Paper [3]). For percutaneous coronary intervention, the survival predictor explained 66% of the hospital level variation in mortality rates; this compares to 48% for observed mortality and 12% for volume of PCI cases. Measures with low reliability or correlation explain little variation. The correlation between the survival predictor and risk-adjusted mortality was (.96) ([16] p. 232), and the amount of variation explained for PCI was 66% [3]. This is a more than adequate level of reliability.

Note: The percentage of hospital level variation in mortality rates explained by the survival predictor is analogous to R squared in regression analysis. [16, p. 228]

	<p>Citations:</p> <p>[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf</p> <p>[21] Morris CN. Parametric Empirical Bayes Inference: Theory and Applications. J Am Stat Assoc 1988;78:47-55.</p> <p>[22] McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) Frontiers in Health Policy Research. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.</p> <p>[23] Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. N Engl J Med. 2003;349:2117-2127.</p> <p>[20] Needleman, J., Buerhaus, P.I., Mattke, S., Stewart, M., and Zelevinsky, M. (2003). Health Services Research 38.6, Part I; 1487-1508.</p> <p>[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. Medical Care, 47(2): 226-233.</p> <p>[25] McGrath, PD., Wennberg, DE., Dickens, Jr., JD., Siewers, AE., Lucas, FL., Malenka, DJ, Kellett, Jr., MA., Ryan, Jr., TJ. (2000) Relation Between Operator and Hospital Volume and Outcomes Following Percutaneous Coronary Interventions in the Era of the Coronary Stent. JAMA, 284(24):3139-3144.</p>
<p>26</p> <p>(2c)</p>	<p>Validity Testing</p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the percutaneous coronary intervention.</p> <p>Analytic Method: We determined the value of our survival predictor (mortality) measure by establishing whether it explained hospital-level variation in risk-adjusted mortality rates and by assessing to what degree it was able to predict future hospital performance. We first estimated the proportion of variation in hospital-level mortality (2000-01) explained by the survival predictor measure using random effects logistic regression models. For these analyses, we estimated the proportional change in the hospital-level variance in mortality rates, which was determined from the standard deviation of the random effect, after adding each measure to the model [14,22].</p> <p>We next compared the ability of the survival predictor measure to the individual measures, mortality rates and hospital volume. We should note that these analyses focus on explaining systematic, or non-random, variation, since measurement error (random error) is accounted for and subtracted from the total variation in all analyses [22,24].</p> <p>We next determined the extent to which the composite measure predicts future risk-adjusted mortality. For this analysis, hospitals were ranked based on each measure from the earlier time period (data from years 2000-01) and divided into four equal size groups (quartiles at the patient level). The subsequent risk-adjusted mortality rates for each quartile of performance were then calculated (data from years 2002-03). We present the subsequent mortality rates across quartiles of the PCI survival predictor measure to graphically demonstrate its usefulness in discriminating among hospitals for the entire spectrum of performance. To compare the predictive ability of the composite measures and individual measures, we also present the subsequent mortality rates in the “worst” compared to the “best” quartile</p>

	<p>in the White Paper ([3], p. 22) "Quartiles of Performance Measures (2000-2001. This table reflects how well the unadjusted survival predictor created on 2000-2001 data compares to risk-adjusted mortality in 2002-2003 data.</p> <p>Note: The risk-adjusted mortality rate for PCI was constructed using standard methods. We determined the ratio of actual deaths or complications to the number of expected deaths (the O/E ratio). The number of expected deaths was the sum over all patients of the predicted probability of death or complications derived from a logistic regression model estimated on all patients undergoing PCI. The dependent variable in the logistic model was death or complications and the independent variables were patient covariates. The patient characteristics included age, gender, race, admission acuity, and co-existing diseases using the Elixhauser method. A zip code level measure of socio-economic status was derived from 2000 census data.</p> <p>Testing Results: While some measures are good at discriminating top performers or bottom performers, this measure is good at prediction across entire spectrum of performance. [See White paper [3]: Figures p. 21-22) for a graphical demonstration of the usefulness of the survival predictor in discriminating among hospitals across the entire spectrum of performance.] To compare the predictive ability of the reliability adjusted survival predictor versus the individual components (volume and observed mortality) we also present the subsequent mortality rates in the "worst" compared to the "best" quartile. In the case of PCI, the Survival Predictor was a better predictor of subsequent risk adjusted mortality than either hospital volume alone or observed mortality alone. [3, p. 20]</p> <p>[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf</p> <p>[22]. McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) Frontiers in Health Policy Research. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.</p> <p>[14] Birkmeyer JD, Dimick JB, Staiger DO. Operative mortality and procedure volume as predictors of subsequent hospital performance. Ann Surg 2006;243:411-417.</p> <p>[24] Zaslavsky AM, Cleary PD. Dimensions of plan performance for sick and healthy members on the Consumer Assessments of Health Plans Study 2.0 survey. Med Care 2002;40:951-964.</p>
<p>27 (2d)</p>	<p>Measure Exclusions <i>Provide evidence to justify exclusion(s) and analysis of impact on measure results during testing.</i></p> <p>Summary of Evidence supporting exclusion(s):</p> <p>Citations for Evidence:</p> <p>Data/sample:</p> <p>Analytic Method:</p> <p>Testing Results:</p>
<p>28 (2e)</p>	<p>Risk Adjustment Testing <i>Summarize the testing used to determine the need (or no need) for risk adjustment and the statistical performance of the risk adjustment method.</i></p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p>

	<p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing PCI.</p> <p>Analytic Method: Sensitivity analysis. We performed a sensitivity analysis to determine whether risk-adjustment of the mortality input was important in improving the predictive ability of the survival predictor measure. Risk-adjustment was performed using logistic regression to estimate expected mortality rates for each hospital based on patient age, gender, race, urgency of operation, median income, and coexisting diseases. Coexisting diseases were determined from secondary diagnostic codes using the methods of Elixhauser (16). The observed mortality rate at each hospital was then divided by the expected mortality rate to yield the ratio of observed/expected deaths (O/E ratio). The O/E ratio was multiplied by the average mortality rate for each operation to yield a risk-adjusted mortality rate. To determine the value of risk-adjustment in the context of selective referral, we compared the ability of risk-adjusted and unadjusted composite measures to predict subsequent performance.</p> <p>Testing Results: In sensitivity analysis, composite measures based on an unadjusted mortality input and a risk-adjusted mortality input had a correlation of (.95) and thus were equally good at predicting future performance (See pages 21-22 in the White Paper [3]).</p> <p>► If outcome or resource use measure not risk adjusted, provide rationale: Because risk-adjusted mortality for PCI is not available publicly except for limited locations, the capacity to use unadjusted mortality is very desirable, especially since it was shown to provide (under this methodology) an equal result. This measure will allow measurement to occur across the United States, providing information to national companies, health plans and consumers.</p>
<p>29 (2g)</p>	<p>Testing comparability of results when more than 1 data method is specified (e.g., administrative claims or chart abstraction) Data/sample: not applicable</p> <p>Analytic Method:</p> <p>Results:</p>
<p>30 (2f)</p>	<p>Provide Measure Results from Testing or Current Use Results from testing</p> <p>Data/sample: same as described above, results for survival predictor in White Paper [3] available on Website and Validation results for composite in [16] Staiger, Dimick et al., Medical Care 2009</p> <p>Methods to identify statistically significant and practically/meaningfully differences in performance: Bayesian Hierarchical methods using new shrinkage estimator Empirical Bayesian methods to determine weights Correlations Calculated the amount of variation predicted by survival predictor as a percentage of all hospital-level variation (adjusted for sampling variation)--analogous to a R-squared from a regression that summarizes the ability of the predictor to explain the hospital level variation in mortality for PCI procedures. Predictor was tested against the "gold standard" --risk adjusted mortality</p> <p>Results: See White Paper [3]</p>
<p>31 (2h)</p>	<p>Identification of Disparities</p> <p>► If measure is stratified by factors related to disparities (i.e. race/ethnicity, primary language, gender, SES, health literacy), provide stratified results:</p> <p>► If disparities have been reported/identified, but measure is not specified to detect disparities, provide rationale: .</p>
<p>USABILITY</p>	

<p>32 (3)</p>	<p>Current Use Testing completed If in use, how widely used Nationally ► If "other," please describe: Survival Predictor for Pancreatectomy and Esophagectomy in use--see URL.</p> <p><input checked="" type="checkbox"/> Used in a public reporting initiative, name of initiative: Leapfrog Hospital Survey Sample report attached <input type="checkbox"/> OR Web page URL: https://www.leapfroggroup.org/cp</p>
<p>33 (3a)</p>	<p>Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)</p> <p>Data/sample:</p> <p>Methods:</p> <p>Results: See following citations reflecting consumer use of mortality information:</p> <p>[10] Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.</p> <p>[11] Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. Quality Review Bulletin 17:242-53.</p> <p>[12] Hibbard, J.H. and Jewett, J.(1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? Medical Care Research and Review., Vol 53(1): 28-47.</p>
<p>34 (3b, 3c)</p>	<p>Relation to other NQF-endorsed™ measures</p> <p>► Is this measure similar or related to measure(s) already endorsed by NQF (on the same topic or the same target population)? <i>Measures can be found at www.qualityforum.org under Core Documents.</i></p> <p><i>Check all that apply</i></p> <p><input type="checkbox"/> Have not looked at other NQF measures <input checked="" type="checkbox"/> Other measure(s) on same topic <input checked="" type="checkbox"/> Other measure(s) for same target population <input type="checkbox"/> No similar or related measures</p> <p>Name and number of similar or related NQF-endorsed™ measure(s): PCI volume measure (NQF# 165) Risk adjusted PCI mortality (NQF# 0136)</p> <p>Are the measure specifications harmonized with existing NQF-endorsed™ measures? Partially harmonized</p> <p>► If not fully harmonized, provide rationale: The PCI Survival Predictor is harmonized with the volume specifications in NQF 165.</p> <p>It is not harmonized with the Risk-adjusted mortality measure--given the data needed for the risk adjustment models is from proprietary registry information (STS/ACC) and that is not available.</p> <p>Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: This measure provides the ability to produce reliable mortality results for low volume hospitals, other measures do not have this capacity. In addition, the access to data nationally for other PCI mortality measures does not exist.</p>
FEASIBILITY	
<p>35 (4a)</p>	<p>How are the required data elements generated? <i>Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data elements are generated concurrent with and as a byproduct of care processes during care delivery (e.g., blood pressure or other assessment recorded by personnel conducting the assessment)</p> <p><input type="checkbox"/> Data elements are generated from a patient survey (e.g., CAHPS)</p> <p><input checked="" type="checkbox"/> Data elements are generated through coding performed by someone other than the person who obtained the original information (e.g., DRG or ICD-9 coding on claims)</p> <p><input checked="" type="checkbox"/> Other, Please describe: Data are currently submitted to Leapfrog via a secure online survey-</p>
<p>36 (4b)</p>	<p>Electronic Sources All data elements</p> <p>► If all data elements are not in electronic sources, specify the near-term path to electronic collection by most providers:</p>

	<p>► Specify the data elements for the electronic health record: volume of PCI procedure, observed death during inpatient stay, following PCI procedure</p>
37 (4c)	<p>Do the specified exclusions require additional data sources beyond what is required for the other specifications? No</p> <p>► If yes, provide justification:</p>
38 (4d)	<p>Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure: It is unlikely that this procedure, or inpatient death will be inaccurately coded or not coded given the high cost of procedure and the accompanying death.</p> <p>Describe how could these potential problems be audited: If problems were identified, a chart review of cases could be performed.</p> <p>Did you audit for these potential problems during testing? No If yes, provide results:</p>
39 (4e)	<p>Testing feasibility Describe what have you learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: Initial results only available for Esophagectomy, Pancreatectomy. CABG will be released in 2009</p>
CONTACT INFORMATION	
40	<p>Web Page URL for Measure Information Describe where users (implementers) should go for more details on specifications of measures, or assistance in implementing the measure. Web page URL: https://leapfrog.medstat.com for access to Survival Predictor White Paper</p>
41	<p>Measure Steward Point of Contact First Name: MI: Last Name: Credentials (MD, MPH, etc.): Organization: The Leapfrog Group % The Academy Street Address: 1150 17th St., NW, Suite 600 City: Washington State: DC ZIP: 20036 Email: Telephone: ext:</p>
42	<p>Measure Developer Point of Contact If different from Measure Steward First Name: Justin MI: B Last Name: Dimick Credentials (MD, MPH, etc.): MD, MPH Organization: Department of Surgery, University of Michigan, M-SCORE offices, Suite 201 and 202 Street Address: 211 N. Fourth Avenue City: Ann Arbor State: MI ZIP: 48104 Email: jdimick@umich.edu Telephone: ext:</p>
ADDITIONAL INFORMATION	
43	<p>Workgroup/Expert Panel involved in measure development Workgroup/panel used</p> <p>► If workgroup used, describe the members' role in measure development: Research team led by Justin Dimick, MD, MPH;</p> <p>► Provide a list of workgroup/panel members' names and organizations:</p> <p>Douglas Staiger Ph.D., Department of Economics and the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth College, Hanover, New Hampshire</p> <p>John D. Birkmeyer, MD Michigan Surgical Collaborative for Outcomes Research and Evaluation Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Onur Baser, Ph.D.</p>

	<p>Michigan Surgical Collaborative for Outcomes Research and Evaluation Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Research supported by the National Institute on Aging</p>
44	<p>Measure Developer/Steward Updates and Ongoing Maintenance Year the measure was first released: 2008 Month and Year of most recent revision: August 2008 What is the frequency for review/update of this measure? Annual When is the next scheduled review/update for this measure? New coefficients for August 2009</p>
45	<p>Copyright statement/disclaimers: none</p>
46	<p>Additional Information: All measure information is available at https://leapfrog.medstat.com Please contact measure developer prior to use to assure all necessary items have been accessed.</p>
47	<p>I have checked that the submission is complete and any blank fields indicate that no information is provided. <input checked="" type="checkbox"/></p>
48	<p>Date of Submission (MM/DD/YY): Revised submission dated 3/18/09</p>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

The measure information you submit will be shared with NQF’s Steering Committees and Technical Advisory Panels to evaluate measures against the NQF criteria of importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Four conditions (as indicated below) must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, *importance to measure and report*, in order to be evaluated against the remaining criteria. References to the specific measure evaluation criteria are provided in parentheses following the item numbers. Please refer to the *Measure Evaluation Criteria* for more information at www.qualityforum.org under Core Documents. Additional guidance is being developed and when available will be posted on the NQF website.

Use the tab or arrow (↓→) keys to move the cursor to the next field (or back ←↑). There are three types of response fields:

- drop-down menus - select one response;
- check boxes - check as many as apply; and
- text fields - you can copy and paste text into these fields or enter text; these fields are not limited in size, but in most cases, we ask that you summarize the requested information.

Please note that URL hyperlinks do not work in the form; you will need to type them into your web browser.

Be sure to answer all questions. Fields that are left blank will be interpreted as no or none. **Information must be provided in this form.** Attachments are not allowed except to provide additional detail or source documents for information that is summarized in this form. If you have important information that is not addressed by the questions, they can be entered into item #46 near the end of the form.

For questions about this form, please contact the NQF Project Director listed in the corresponding call for measures.

CONDITIONS FOR CONSIDERATION BY NQF	
	Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards.
A (A)	Public domain or Measure Steward Agreement signed: Public domain - Agreement not required (If no, do not submit) <i>Template for the Measure Steward Agreement is available at www.qualityforum.org under Core Documents.</i>
B (B)	Measure steward/maintenance: Is there an identified responsible entity and process to maintain and update the measure on a schedule commensurate with clinical innovation, but at least every 3 years? Yes, information provided in contact section (If no, do not submit)
C (C)	Intended use: Does the intended use of the measure include BOTH public reporting AND quality improvement? Yes (If no, do not submit)
D (D)	Fully developed and tested: Is the measure fully developed AND tested? Yes, fully developed and tested (If not tested and no plans for testing within 24 months, do not submit)

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

	(for NQF staff use) NQF Review #: HOE-021-08 NQF Project: Hospital Outcomes and Efficiency
MEASURE SPECIFICATIONS & DESCRIPTIVE INFORMATION	
1	Information current as of (date- MM/DD/YY):
2	Title of Measure: Survival Predictor for Abdominal Aortic Aneurysm (AAA)
3	Brief description of measure ¹ : A reliability adjusted measure of AAA repair performance that optimally combines two important domains: AAA hospital volume and AAA operative mortality, to provide predictions on AAA survival rates for hospitals. This measure is calculated based on data from administrative claims information.
4 (2a)	<p>Numerator Statement: Note: Because of the type of modeling done for this Survival Predictor--the information is not readily split into Numerator/ Denominator statements. Thus, we describe the two domains and their coding and data needs in this section.</p> <p>The formula for calculating the survival predictor has two components, one is a volume predicted mortality rate, and the second is an observed mortality rate. The volume predicted mortality rate reflects the hospitals experience performing AAA surgeries (thus, it includes all AAA surgeries) and uses mortality for all hospitals at that specific volume to create the volume predicted mortality. The input data from the hospitals for this domain is a volume count of all AAAs performed in the hospital.</p> <p>The second domain is the observed mortality, for this domain the population is the group of AAA cases, the data needed for this domain is the number of observed deaths occurring for AAA cases, within the inpatient setting.</p> <p>Note: All data is available in administrative claims information. In the case of Leapfrog's implementation hospitals are asked to submit aggregated information from their claims data. No personal health information is submitted to Leapfrog. Other users of the measure may have direct access to administrative data.</p> <p>Time Window: Annual</p> <p>Numerator Details (Definitions, codes with description): For the volume predicted mortality, hospitals count the number of AAA cases using the following codes:</p> <p>ICD-9-CM Procedure</p> <ul style="list-style-type: none"> ■ 3834 Aorta Resection & Anast ■ 3844 Resection Abdominal Aorta with replacement ■ 3864 Excision of aorta ■ 3925 Aorta-iliac-femoral bypass ■ 3971 Endo Implant of Graft in Aorta <p>Exclude:</p> <ul style="list-style-type: none"> ■ 3845 thoracoabdominal procedures <p>See calculation worksheet for details on how volume-predicted mortality is used in the model.</p>

¹ Example of measure description: Percentage of adult patients with diabetes aged 18-75 years receiving one or more A1c test(s) per year.
NQF Measure Submission Form, V3.1

	<p>For the observed mortality domain, the hospital submits the observed deaths for AAA cases without rupture using the following codes:</p> <ul style="list-style-type: none"> ■ 3834 Aorta Resection & Anast ■ 3844 Resection Abdominal Aorta with replacement ■ 3864 Excision of aorta ■ 3925 Aorta-iliac-femoral bypass ■ 3971 Endo Implant of Graft in Aorta <p>and includes Diagnosis Codes for:</p> <p>441.4 Dissection of aorta aneurysm unspecified site 441.7 Thoracoabdominal aneurysm without rupture 441.9 Aortic aneurysm of unspecified site without rupture</p> <p>Mortality Domain excludes all Thoracic Diagnosis Codes and dissection codes for AAA without Rupture</p> <p>441.0x General code 441.1 Thoracic aneurysm ruptured 441.2 Thoracic aneurysm without rupture 441.3 Abdominal aneurysm ruptured 441.5 Aortic aneurysm of unspecified site ruptured 441.6 Thoracoabdominal aneurysm ruptured</p> <p>Mortality Domain does excludes thoracic aneurysm Procedure Code:</p> <p>38.45 Resection of vessel with replacement, other thoracic vessels</p> <p>See Calculation Worksheet for examples of how the two domains are used to create the Survival Predictor.</p>
<p>5 (2a)</p>	<p>Denominator Statement:</p> <p>Time Window:</p> <p>Denominator Details (Definitions, codes with description):</p>
<p>6 (2a, 2d)</p>	<p>Denominator Exclusions: No exclusions)</p> <p>Denominator Exclusion Details (Definitions, codes with description): (</p>
<p>7 (2a, 2h)</p>	<p>Stratification Do the measure specifications require the results to be stratified? No</p> <p>► If "other" describe:</p> <p>Identification of stratification variable(s):</p> <p>Stratification Details (Definitions, codes with description):</p>
<p>8 (2a, 2e)</p>	<p>Risk Adjustment Does the measure require risk adjustment to account for differences in patient severity before the onset of care? No ► If yes, (select one)</p> <p>► Is there a separate proprietary owner of the risk model? No</p> <p>Identify Risk Adjustment Variables: See section 28 for rationale and support for not risk adjusting this measure. Measure was tested against risk adjusted mortality--details on that provided in Section 26.</p> <p>Detailed risk model: attached <input type="checkbox"/> OR Web page URL:</p>
<p>9</p>	<p>Type of Score: Rate/proportion Calculation Algorithm: attached <input checked="" type="checkbox"/> OR Web page URL:</p>

(2a)	<p>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 = Score within a defined interval ▶ If "Other", please describe:</p>																
<p>10 (2a, 4a, 4b)</p>	<p>Identify the required data elements(e.g., primary diagnosis, lab values, vital signs): procedure codes Data dictionary/code table attached <input checked="" type="checkbox"/> OR Web page URL: Data Quality (2a) Check all that apply <input checked="" type="checkbox"/> Data are captured from an authoritative/accurate source (e.g., lab values from laboratory personnel) <input checked="" type="checkbox"/> Data are coded using recognized data standards <input checked="" type="checkbox"/> Method of capturing data electronically fits the workflow of the authoritative source <input type="checkbox"/> Data are available in EHRs <input checked="" type="checkbox"/> Data are auditable</p>																
<p>11 (2a, 4b)</p>	<p>Data Source and Data Collection Methods Identifies the data source(s) necessary to implement the measure specifications. Check all that apply</p> <table border="0"> <tr> <td><input type="checkbox"/> Electronic Health/Medical Record</td> <td><input type="checkbox"/> Paper Medical Record</td> </tr> <tr> <td><input type="checkbox"/> Electronic Clinical Database, Name:</td> <td><input type="checkbox"/> Standardized clinical instrument, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Clinical Registry, Name:</td> <td><input type="checkbox"/> Standardized patient survey, Name:</td> </tr> <tr> <td><input checked="" type="checkbox"/> Electronic Claims</td> <td><input type="checkbox"/> Standardized clinician survey, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Pharmacy data</td> <td><input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.</td> </tr> <tr> <td><input type="checkbox"/> Electronic Lab data</td> <td></td> </tr> <tr> <td><input type="checkbox"/> Electronic source - other, Describe:</td> <td></td> </tr> </table> <p style="text-align: right;">Instrument/survey attached <input checked="" type="checkbox"/> OR Web page URL:</p>	<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record	<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:	<input type="checkbox"/> Electronic Clinical Registry, Name:	<input type="checkbox"/> Standardized patient survey, Name:	<input checked="" type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:	<input type="checkbox"/> Electronic Pharmacy data	<input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.	<input type="checkbox"/> Electronic Lab data		<input type="checkbox"/> Electronic source - other, Describe:			
<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record																
<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:																
<input type="checkbox"/> Electronic Clinical Registry, Name:	<input type="checkbox"/> Standardized patient survey, Name:																
<input checked="" type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:																
<input type="checkbox"/> Electronic Pharmacy data	<input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.																
<input type="checkbox"/> Electronic Lab data																	
<input type="checkbox"/> Electronic source - other, Describe:																	
<p>12 (2a)</p>	<p>Sampling If measure is based on a sample, provide instructions and guidance on sample size. Minimum sample size: Instructions:</p>																
<p>13 (2a)</p>	<p>Type of Measure: Outcome ▶ If "Other", please describe: ▶ If part of a composite or paired with another measure, please identify composite or paired measure While the measure uses two types of information components (domains), the results are not a composite as is defined by NQF, but rather a reliability adjusted measure of survival. Volume is used to create a volume predicted mortality for the hospital--this component of the measure is used to create greater reliability for low-volume hospitals. In the modeling for this measure, the volume predicted mortality and the observed mortality are weighted. In the model, lower volume hospitals have a higher weight on the volume predicted mortality versus the observed mortality. The opposite is true for high volume hospitals, which have a higher weight on the observed mortality. This methodology results in a reliability adjusted survival predictor.</p>																
<p>14 (2a)</p>	<p>Unit of Measurement/Analysis (Who or what is being measured) Check all that apply.</p> <table border="0"> <tr> <td><input type="checkbox"/> Can be measured at all levels</td> <td><input type="checkbox"/> Integrated delivery system</td> </tr> <tr> <td><input type="checkbox"/> Individual clinician (e.g., physician, nurse)</td> <td><input checked="" type="checkbox"/> Health plan</td> </tr> <tr> <td><input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)</td> <td><input checked="" type="checkbox"/> Community/Population</td> </tr> <tr> <td><input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)</td> <td><input type="checkbox"/> Other (Please describe):</td> </tr> </table>	<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system	<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input checked="" type="checkbox"/> Health plan	<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input checked="" type="checkbox"/> Community/Population	<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other (Please describe):								
<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system																
<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input checked="" type="checkbox"/> Health plan																
<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input checked="" type="checkbox"/> Community/Population																
<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other (Please describe):																
<p>15 (2a)</p>	<p>Applicable Care Settings Check all that apply</p> <table border="0"> <tr> <td><input type="checkbox"/> Can be used in all healthcare settings</td> <td><input type="checkbox"/> Hospice</td> </tr> <tr> <td><input type="checkbox"/> Ambulatory Care (office/clinic)</td> <td><input checked="" type="checkbox"/> Hospital</td> </tr> <tr> <td><input type="checkbox"/> Behavioral Healthcare</td> <td><input type="checkbox"/> Long term acute care hospital</td> </tr> <tr> <td><input type="checkbox"/> Community Healthcare</td> <td><input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)</td> </tr> <tr> <td><input type="checkbox"/> Dialysis Facility</td> <td><input type="checkbox"/> Prescription Drug Plan</td> </tr> <tr> <td><input type="checkbox"/> Emergency Department</td> <td><input type="checkbox"/> Rehabilitation Facility</td> </tr> <tr> <td><input type="checkbox"/> EMS emergency medical services</td> <td><input type="checkbox"/> Substance Use Treatment Program/Center</td> </tr> <tr> <td><input type="checkbox"/> Health Plan</td> <td><input type="checkbox"/> Other (Please describe):</td> </tr> </table>	<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice	<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital	<input type="checkbox"/> Behavioral Healthcare	<input type="checkbox"/> Long term acute care hospital	<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)	<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan	<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility	<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center	<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other (Please describe):
<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice																
<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital																
<input type="checkbox"/> Behavioral Healthcare	<input type="checkbox"/> Long term acute care hospital																
<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)																
<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan																
<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility																
<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center																
<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other (Please describe):																

	<input type="checkbox"/> Home Health
IMPORTANCE TO MEASURE AND REPORT	
Note: This is a threshold criterion. If a measure is not judged to be sufficiently important to measure and report, it will not be evaluated against the remaining criteria.	
16 (1a)	<p>Is measure related to a National Priority Partners priority area? Safety reliability (for NQF staff use) Does measure address a <u>specific</u> NPP goal? (www.qualityforum.org/about/NPP/):</p>
17 (1a)	<p>Does the measure address a high impact aspect of healthcare patient/societal consequences of poor quality</p> <p>Summary of Evidence: This measure addresses mortality in a high risk procedure (AAA) and is an outcome measure which is of interest to both consumers and purchasers.</p> <p>As derived from the National Inpatient Sample (NIS), 48% of patients currently undergo AAA repair at hospitals performing fewer than 50 procedures per year. Adjusted mortality rates were significantly higher at such hospitals (5.1%) than at hospitals exceeding Leapfrog's volume criteria (3.8%) [1c].</p> <p>In addition to addressing high volume procedure risk, this measure improves upon the technology of surgical procedure mortality measurement. It overcomes three problems with existing AAA mortality measures: 1) Mortality rates are often too "noisy" to reflect hospital quality with surgery (particularly among lower volume hospitals), 2) volume alone is a weak proxy for most procedures, and 3) when both volume and mortality are reported as separate indicators it is difficult to understand which measure is more important. [1] Given that 48% of patients have AAA procedures performed at low volume hospitals in the United States, and that this measure specifically addresses hospitals which perform elective procedures, consumers and purchasers would benefit from information that is more reliable in the prediction of future mortality for both selection and selective referral. In addition, this measure can be applied to the nation, states, or regions.</p> <p>Birkmeyer and Dimick (2009)[4] show that differences in mortality can be predicted using a reliability adjusted mortality rate (a weighted combination of volume and mortality) which is particularly relevant for selective-referral or public reporting contexts. They reduce the effects of random chance (statistical noise) and as a result with CABG, for example, more than half of the observed variation can be attributed to statistical noise. When they sorted hospitals simply on actual (risk-adjusted) mortality, rates varied from 1.4% to 11.0% across hospital quintiles (Figure 1 in White Paper [3]). After they adjusted for reliability, however, the mortality rates varied considerably less, from 3.3% to 6.3%.</p> <p>Citations² for Evidence:</p> <p>[1a] Rathore, A.J., Epstein, A.J., Rathore, S.S., Volpp, K.G., and Krunholz, H.M. (2004). Hospital percutaneous coronary intervention volume and patient mortality. 1988 to 2000: does the evidence support current procedure volume minimums? <i>J Am Coll Surg.</i>; 43:(10): 1755-62.</p> <p>[1b] Moscucci, M., Eagle, K.A., Share, D., Smith, D., DeFranco, A.C., O'Donnell, M., Kline-Rogers, E., Jani, S.M., and Brown, D.L. (2005). Public Reporting and Case Selection for Percutaneous Coronary Interventions: An Analysis from Two Large Multicenter Percutaneous Coronary Intervention Databases. <i>J. Am Coll Card.</i>, 45(11):1759-1765.</p> <p>[1c] Birkmeyer, JD., and Dimick JB. The Leapfrog Group's Safety Practices, 2003: The Potential Benefits of Universal Adoption. Available on the Leapfrog Group website: www.leapfroggroup.org</p> <p>[2] The National Hospital Bill: The Most Expensive Conditions by Payer, 2006. Statistical Brief #59. File accessed on March 16, 2009, at: http://www.hcup-us.ahrq.gov/reports/statbriefs/sb59.jsp Produced by</p>

² Citations can include, but are not limited to journal articles, reports, web pages (URLs).
NQF Measure Submission Form, V3.1

	<p>AHRQ, Center for Delivery, Organization, and Markets, Healthcare Cost and Utilization Project, Nationwide Inpatient Sample, 2006.</p> <p>[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf</p> <p>[4] Birkmeyer, J.D., and Dimick, J.B. (2009) Understanding and reducing variation in surgical mortality. <i>Annu. Rev. Med.</i> 2009. 60:405-15.</p> <p>[4b] Birkmeyer, J.D., Siewers, A.E., Finlayson, E.V.A, Stukel, T.A., Lucas, F.L., Batista, I., Welch, G., Wennberg, D.A.. (2002) Hospital Volume and Surgical Mortality in US. <i>N Engl J Med</i>, Vol. 346, No. 15 •1128-1137.</p>
<p>18 (1b)</p>	<p>Opportunity for Improvement <i>Provide evidence that demonstrates considerable variation, or overall poor performance, across providers.</i></p> <p>Summary of Evidence: In 2002, a systematic review of the literature on the volume-outcome relationship found that there was a significant relationship between hospital volume and outcomes. [5]</p> <p>The absolute differences in adjusted mortality rates between very-low volume hospitals and very-high-volume hospitals were slightly more than 1 percent for AAA repair. [1c]</p> <p>Given the findings related to volume of procedures, Silber et al., [7] explored the relative contribution of complication rates and failure to rescue rates to mortality and found that complication rates were more likely influenced by patient factors while failure to rescue rates of those with complications was more related to hospital factors. Thus, it may be that higher volume hospitals are better at rescuing patients with complications. Silbers finding, in conjunction with the volume information, suggests lower volume hospitals with worse mortality rates could in fact address this through better care following the procedure, thereby reducing their overall rate. Unfortunately, most low volume hospitals in the United States do not have information on their AVRI mortality rate compared to other hospitals. When they are given this information, there is a good chance for improvement. Birkmeyer and Dimick [4] report that in northern New England, mortality associated with CABG fell by >25% when hospitals and surgeons were given feedback on their mortality data. We expect that a similar reaction would occur for AAA repair.</p> <p>Note: Studies [4, 6] indicate it is also likely that some lower volume hospitals would also have lower mortality rates.</p> <p>Citations for Evidence:</p> <p>[1a] Rathore, A.J., Epstein, A.J., Rathore, S.S., Volpp, K.G., and Krunholz, H.M. (2004). Hospital percutaneous coronary intervention volume and patient mortality. 1988 to 2000: does the evidence support current procedure volume minimums? <i>J Am Coll Surg.</i>; 43:(10): 1755-62.</p> <p>[1c] Birkmeyer, JD., and Dimick JB. The Leapfrog Group's Safety Practices, 2003: The Potential Benefits of Universal Adoption. Available on the Leapfrog Group website: www.leapfroggroup.org</p> <p>[4b] Birkmeyer, J.D., Siewers, A.E., Finlayson, E.V.A, Stukel, T.A., Lucas, F.L., Batista, I., Welch, G., Wennberg, D.A.. (2002) Hospital Volume and Surgical Mortality in US. <i>N Engl J Med</i>, Vol. 346, No. 15 •1128-1137.</p> <p>[5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. <i>Annals of Internal Medicine</i>, Sept 1;137(6):511-20</p> <p>[6] Epstein, A.J., Rathore, S.S., Krumholz, H.M., and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p>

	<p>[7] Silber, J.H., Rosenbaum, P.R., Trudeau, M.E., et al. 2005. Changes in prognosis after the first postoperative complication. <i>Medical Care</i>, 43:122-31.</p>
<p>19 (1b)</p>	<p>Disparities <i>Provide evidence that demonstrates disparity in care/outcomes related to the measure focus among populations.</i></p> <p>Summary of Evidence: It is more likely that minorities will be treated at a low volume facility, and as a result are likely to be impacted by higher mortality rates. In an analysis of the National Inpatient Sample, Epstein, Rathore and Krumholz (2005)[6, pags 3-5] found that a greater proportion of patients treated in low volume hospitals for both CABG and PCI conditions were non-white, while a lower proportion of non-white patients presented as "elective" admissions or patients received in transfer as compared to patients in high volume hospitals. We expect that is would be similar for AAA.</p> <p>Citations for evidence:</p> <p>[6] Epstein, A.J., Rathore, S.S., Krumholz, H.M., and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p>
<p>20 (1c)</p>	<p>If measuring an Outcome Describe relevance to the national health goal/priority, condition, population, and/or care being addressed: AAA is a high risk surgery, with mortality rates with 5% mortality rates in half of the hospitals performing AAA procedures. [1c]</p> <p>This measure is highly relevant to both consumers and purchasers, given its frequency. National purchasers are interested in comparative information on hospitals nationwide. Pauly (1996) in a study of purchaser interests in hospital performance reporting found that mortality ratings were more important to purchasers than were morbidity or complications. [9] Health plans are interested in contracting with centers of excellence, which can be identified through the results of survival predictor in combination with other information on cost and quality. Consumers have shown their interest in cardiac procedure mortality by requesting reports from the state of Pennsylvania [10]; an earlier study by IOM (Lohr, Donaldson and Walker 1991) found that consumers were interested in hospital mortality rates, but did not perceive this information to be available.[11] Hibbard and Jewett found that consumers were more interested in "undesirable events" (such as mortality, complications, infections) than in "desirable events." [12]</p> <p>[9] Pauly, M.V., Brailer, D.J.Kroch, E., and Even-Shoshan, O. Measuring Hospital Outcomes from a Buyer's Perspective. <i>American Journal of Medical Quality</i>, 11(8): Fall 1996.</p> <p>[10]Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.</p> <p>[11]Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. <i>Quality Review Bulletin</i> 17:242-53.</p> <p>[12]Hibbard, J.H. and Jewett, J.(1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? <i>Medical Care Research and Review.</i>, Vol 53(1): 28-47.</p> <p>If not measuring an outcome, provide evidence supporting this measure topic and grade the strength of the evidence <i>Summarize the evidence (including citations to source) supporting the focus of the measure as follows:</i></p> <ul style="list-style-type: none"> • <u>Intermediate outcome</u> - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit. • <u>Process</u> - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and

if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).

- **Structure** - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
- **Patient experience** - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
- **Access** - evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
- **Efficiency**- demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

Type of Evidence *Check all that apply*

<input type="checkbox"/> Evidence-based guideline	<input checked="" type="checkbox"/> Quantitative research studies
<input checked="" type="checkbox"/> Meta-analysis	<input type="checkbox"/> Qualitative research studies
<input checked="" type="checkbox"/> Systematic synthesis of research	<input type="checkbox"/> Other (<i>Please describe</i>):

Overall Grade for Strength of the Evidence³ (*Use the USPSTF system, or if different, also describe how it relates to the USPSTF system*): **Moderate**

Summary of Evidence (*provide guideline information below*): **Over 100 articles published related to volume and outcome relationship, with some inconsistency in results. Systematic review of the literature conducted in 2002. No review since that time.**

Citations for Evidence:

[1c] Birkmeyer, JD., and Dimick JB. The Leapfrog Group's Safety Practices, 2003: The Potential Benefits of Universal Adoption. Available on the Leapfrog Group website: www.leapfroggroup.org

[5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. *Annals of Internal Medicine*, Sept 1;137(6):511-20.

[14] Birkmeyer, J.D., Dimick, J.B., Staiger, D.O. (2006) Operative mortality and procedure volume as predictors of subsequent hospital performance. *Ann Surg.* 243:411-417.

[15] Dimick, JB, Welch HG, Birkmeyer JD. (2004) Surgical mortality as an indicator of hospital quality: The problem with small sample size. *JAMA*, 292:847-851.

[4] Birkmeyer, JD., and Dimick, JB. (2009) Understanding and reducing variation in surgical mortality. *Annu. Rev. Med.* 60:405-15.

[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. *Medical Care*, 47(2): 226-233.

[6] Andrew J Epstein, Saif S Rathore, Harlan M Krumholz and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. *BMC Health Services Research*, 2005, vol. 5:42, accessed at: <http://www.biomedcentral.com/1472-6963/5/42>

18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. *N Engl J Med.* 1979;301:1364-9.

³The strength of the body of evidence for the specific measure focus should be systematically assessed and rated, e.g., USPSTF grading system www.ahrq.gov/clinic/uspstmeth.htm: **A** - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. **B** - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. **C** - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. **D** - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. **I** - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

<p>21 (1c)</p>	<p>Clinical Practice Guideline <i>Cite the guideline reference; quote the specific guideline recommendation related to the measure and the guideline author's assessment of the strength of the evidence; and summarize the rationale for using this guideline over others.</i></p> <p>Guideline Citation:</p> <p>Specific guideline recommendation:</p> <p>Guideline author's rating of strength of evidence <i>(If different from USPSTF, also describe it and how it relates to USPSTF):</i></p> <p>Rationale for using this guideline over others: .</p>
<p>22 (1c)</p>	<p>Controversy/Contradictory Evidence <i>Summarize any areas of controversy, contradictory evidence, or contradictory guidelines and provide citations.</i></p> <p>Summary: There are three areas of possible contention with this measure--</p> <p>1) The volume-outcome relationship has been questioned for some procedures [6, 17, 19, 25]</p> <p>Epstein and Rathore et al., [6] questioned whether it was appropriate to move patients from low volume hospitals to high volume hospitals, given the number of patients that would have to be moved to save 1 life. They did find in their study of CABG and PCI's performed in the US that low volume hospitals did have higher unadjusted and adjusted for case mix mortality. Of concern, is that 38% of all CABG surgery is performed in low volume hospitals; and that non-white patients were more likely to be treated at low volume hospitals.</p> <p>Peterson et al., [19] questioned the volume outcomes relationship for CABG surgery, and found only modest associations for volume and outcome for CABG. Those with high volume had mortality rate of 2.5% while low volume hospitals rate was 3.2%. They suggest using past mortality rate to select hospitals. (The survival predictor uses both volume and mortality to predict survival in the next year.)</p> <p>Yet, more than 100 studies have demonstrated better results at high-volume hospitals with cardiovascular surgery, major cancer resections, and other high-risk procedures.[18, 20]</p> <p>There is specific evidence of the variation in mortality across the different volume levels of AAAs. {1c}. They documented that there were differences between low volume hospitals and high volume hospitals in mortality--with high volume hospitals having less mortality.</p> <p>2) That outcome measures must be risk-adjusted unless there is evidence to show it is not needed (NQF). The survival predictor measure predicts better than volume or mortality alone, and is as good a predictor as risk-adjusted mortality. When testing the unadjusted survival predictor against risk-adjusted mortality there was a (.96) correlation. [4] See Section 28 of this form for details.</p> <p>3) The weighting of input measures into composites. Existing approaches rely on overly simplistic approaches. Among these, assigning equal weight to all measures (i.e., the all or none approach) and relying on expert opinion are the most common. The survival predictor relies on empiric methods for weighting the input measures.</p> <p>Citations:</p> <p>[1a] Rathore, A.J., Epstein, A.J., Rathore, S.S., Volpp, K.G., and Krumholz, H.M. (2004). Hospital percutaneous coronary intervention volume and patient mortality. 1988 to 2000: does the evidence support current procedure volume minimums? <i>J Am Coll Surg.</i>; 43:(10): 1755-62.</p> <p>[6] Andrew J Epstein, Saif S Rathore, Harlan M Krumholz and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p>

	<p>[16]Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. <i>Medical Care</i>, 47(2): p. 232.</p> <p>[17] Edward L. Hannan, PhD; Chuntao Wu, PhD; Thomas J. Ryan, MD; Edward Bennett, MD; Alfred T. Culliford, MD; Jeffrey P. Gold, MD; Alan Hartman, MD; O. Wayne Isom, MD; Robert H. Jones, MD; Barbara McNeil, MD, PhD; Eric A. Rose, MD; Valavanur A. Subramanian, MD. Do Hospitals and Surgeons With Higher Coronary Artery Bypass Graft Surgery Volumes Still Have Lower Risk-Adjusted Mortality Rates? <i>Circulation</i>. 2003;108:795-801.</p> <p>[18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. <i>N Engl J Med</i>. 1979;301:1364-9.</p> <p>[19] Eric D. Peterson, MD, MPH; Laura P. Coombs, PhD; Elizabeth R. DeLong, PhD; Constance K. Haan, MD; T. Bruce Ferguson, MD. Procedural Volume as a Marker of Quality for CABG Surgery. <i>JAMA</i>. 2004;291:195-201.</p> <p>[20] Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. <i>JAMA</i>. 1998;280:1747-51.)</p> <p>[25] McGrath, PD., Wennberg, DE., Dickens, Jr., JD., Siewers, AE., Lucas, FL., Malenka, DJ, Kellett, Jr., MA., Ryan, Jr., TJ. (2000) Relation Between Operator and Hospital Volume and Outcomes Following Percutaneous Coronary Interventions in the Era of the Coronary Stent. <i>JAMA</i>, 284(24):3139-3144.</p>
<p>23 (1)</p>	<p>Briefly describe how this measure (as specified) will facilitate significant gains in healthcare quality related to the specific priority goals and quality problems identified above: This measure of predicted survival improves upon the reliability of mortality results for high risk surgical procedures, such as AAA. For the first time, this measure produces reliable mortality/survivability information on smaller volume hospitals, as well as high volume hospitals. Hospitals across the country will have information available through voluntary public reporting.</p>
<p>SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES</p>	
<p>Note: Testing and results should be summarized in this form. However, additional detail and reports may be submitted as supplemental information or provided as a web page URL. If a measure has not been tested, it is only potentially eligible for time-limited endorsement.</p>	
<p>24</p>	<p>Supplemental Testing Information: attached <input checked="" type="checkbox"/> OR Web page URL:</p>
<p>25 (2b)</p>	<p>Reliability Testing</p> <p>Data/sample: Data was a 100% sample from the Medicare Analysis Provider and Review (MEDPAR) files for 2000-2003, these files contain 100% of Medicare hospitalizations for years specified. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing an Abdominal aortic aneurysm repair. Codes used to select patients are indicated in Section 4 of this form.</p> <p>This dataset has been used by other researchers to look at the issue of volume and mortality for PCI. As indicated by McGrath and Wennberg et al., [25] the Medicare dataset allowed sufficient power to determine significant differences in adverse outcomes across varying levels of volume.</p> <p>Note: Needleman, Buerhaus, et al., (2003) concluded after applying operational tests on Medicare data for adverse outcomes and all-patient hospital data from 11 states, that Medicare data could be used to assess quality in hospitals.[20] Given the lack of a national all-patient/all-payer database, MEDPAR data</p>

was used in development and testing of the models.

Analytic Method:

Model Development

We used an empirical Bayes approach to combine mortality rates with information on hospital volume at each hospital. In traditional empirical Bayes methods, a point estimate (e.g., mortality rate observed at a hospital) is adjusted for reliability by shrinking it towards the overall mean (e.g., overall mortality rate in the population) [21,22]. We modified this traditional approach by shrinking the observed mortality rate back toward the mortality rate expected given the volume at that hospital—we refer to this as the “volume-predicted mortality” (See attached White Paper [3] TECHNICAL APPENDIX for the mathematical details of this method). With this approach, the observed mortality rate is weighted according to how reliably it is estimated, with the remaining weight placed on the information regarding hospital volume. Because this method includes observed data to the extent that it is useful, and only relies on the proxy measure to the extent necessary, it ensures an optimal combination of these two quality domains. [3]

The two inputs to the survival predictor measure are mortality rates and procedure volume for each of the six included operations. Procedure-specific mortality rates were calculated for all hospitals over a 2-year period (2000-01) and this was used as the first input. Hospital volume was calculated as the number of Medicare cases performed during the same time period. For each operation, the relationship between hospital volume and risk-adjusted mortality was modeled using linear regression. (Details of the risk-adjustment strategy will be discussed below.) After testing the fit of several transformations, hospital volume was modeled as the natural log of the continuous volume variable, which is the same approach used in our previous work [23]. Using this regression model, we estimated the volume-predicted mortality, the second input to the survival predictor measure.

We then used the empirical Bayes approach to create an optimal combination of these two inputs. This survival predictor measure theoretically provides the best estimate of a hospital's true mortality rate, taking into account the both available inputs [21,22]. The combined survival predictor measure was calculated as follows: mortality prediction = (weight)*(observed mortality) + (1-weight)*(volume-predicted mortality). The weight placed on the point estimate of mortality is the reliability, or ratio of signal to signal plus noise, calculated as follows: weight = variation among hospitals/(variation among hospitals + variation within hospitals). The variation among hospitals was calculated as the variance in observed mortality rates for the hospitals included in the sample. The variation within hospitals was calculated as the standard error of the mortality rate at each hospital. With this method, more weight is placed on the observed mortality rate when a hospital has a high number of cases because it is estimated with more reliability; less weight is placed on the observed mortality rate when a hospital performs a low number of cases because of its lower reliability.

A calculation worksheet with examples is attached.

Testing Results: Hospital caseloads and the weights applied to each input to the survival predictor measure varied for each procedure studied (see Table 1 in white paper [3]). For abdominal aortic aneurysm, a procedure with lower hospital caseloads than CABG or PCI, the weight applied to the volume input was .71. ([3]-Table 1). For hospitals with higher volumes more weight was placed on the observed mortality.

The survival predictor (mortality) measure explained a large proportion of non-random, hospital-level variation in risk-adjusted mortality rates (see Table 2, p. 19 in White Paper [3]). For abdominal aortic aneurysm repair the survival predictor explained 41% of the hospital level variation in mortality rates; this compares to 21% for observed mortality and 28% for volume of AAA cases. Measures with low reliability or correlation explain little variation. The correlation between the survival predictor and risk-adjusted mortality was (.96) ([16] p. 232), and the amount of variation explained for elective AAA was 41% [3]. This is a more than adequate level of reliability.

	<p>Note: The percentage of hospital level variation in mortality rates explained by the survival predictor is analogous to R squared in regression analysis. [16, p. 228]</p> <p>Citations:</p> <p>[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf</p> <p>[21] Morris CN. Parametric Empirical Bayes Inference: Theory and Applications. J Am Stat Assoc 1988;78:47-55.</p> <p>[22] McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) Frontiers in Health Policy Research. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.</p> <p>[23] Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. N Engl J Med. 2003;349:2117-2127.</p> <p>[20] Needleman, J., Buerhaus, P.I., Mattke, S., Stewart, M., and Zelevinsky, M. (2003). Health Services Research 38.6, Part I; 1487-1508.</p> <p>[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. Medical Care, 47(2): 226-233.</p> <p>[25] McGrath, PD., Wennberg, DE., Dickens, Jr., JD., Siewers, AE., Lucas, FL., Malenka, DJ, Kellett, Jr., MA., Ryan, Jr., TJ. (2000) Relation Between Operator and Hospital Volume and Outcomes Following Percutaneous Coronary Interventions in the Era of the Coronary Stent. JAMA, 284(24):3139-3144.</p>
<p>26</p> <p>(2c)</p>	<p>Validity Testing</p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the elective abdominal aortic aneurysm repair surgery.</p> <p>Analytic Method: We determined the value of our survival predictor (mortality) measure by establishing whether it explained hospital-level variation in risk-adjusted mortality rates and by assessing to what degree it was able to predict future hospital performance. We first estimated the proportion of variation in hospital-level mortality (2000-01) explained by the survival predictor measure using random effects logistic regression models. For these analyses, we estimated the proportional change in the hospital-level variance in mortality rates, which was determined from the standard deviation of the random effect, after adding each measure to the model [14,22].</p> <p>We next compared the ability of the survival predictor measure to the individual measures, mortality rates and hospital volume. We should note that these analyses focus on explaining systematic, or non-random, variation, since measurement error (random error) is accounted for and subtracted from the total variation in all analyses [22,24].</p> <p>We next determined the extent to which the composite measure predicts future risk-adjusted mortality. For this analysis, hospitals were ranked based on each measure from the earlier time period (data from years 2000-01) and divided into four equal size groups (quartiles at the patient level). The subsequent risk-adjusted mortality rates for each quartile of performance were then calculated (data from years 2002-03). We present the subsequent mortality rates across quartiles of the AVR survival predictor</p>

	<p>measure to graphically demonstrate its usefulness in discriminating among hospitals for the entire spectrum of performance. To compare the predictive ability of the composite measures and individual measures, we also present the subsequent mortality rates in the “worst” compared to the “best” quartile in the White Paper ([3], p. 22) “Quartiles of Performance Measures (2000-2001). This table reflects how well the unadjusted survival predictor created on 2000-2001 data compares to risk-adjusted mortality in 2002-2003 data.</p> <p>Note: The risk-adjusted mortality rate for AAA was constructed using standard methods. We determined the ratio of actual deaths or complications to the number of expected deaths (the O/E ratio). The number of expected deaths was the sum over all patients of the predicted probability of death or complications derived from a logistic regression model estimated on all patients undergoing AAAI. The dependent variable in the logistic model was death or complications and the independent variables were patient covariates. The patient characteristics included age, gender, race, admission acuity, and co-existing diseases using the Elixhauser method. A zip code level measure of socio-economic status was derived from 2000 census data.</p> <p>Testing Results: While some measures are good at discriminating top performers or bottom performers, this measure is good at prediction across entire spectrum of performance. [See White paper [3]: Figures p. 21-22) for a graphical demonstration of the usefulness of the survival predictor in discriminating among hospitals across the entire spectrum of performance.] To compare the predictive ability of the reliability adjusted survival predictor versus the individual components (volume and observed mortality) we also present the subsequent mortality rates in the “worst” compared to the “best” quartile. In the case of AAA, the Survival Predictor was a better predictor of subsequent risk adjusted mortality than either hospital volume alone or observed mortality alone. [3, p. 20]</p> <p>[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf</p> <p>[22]. McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) Frontiers in Health Policy Research. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.</p> <p>[14] Birkmeyer JD, Dimick JB, Staiger DO. Operative mortality and procedure volume as predictors of subsequent hospital performance. Ann Surg 2006;243:411-417.</p> <p>[24] Zaslavsky AM, Cleary PD. Dimensions of plan performance for sick and healthy members on the Consumer Assessments of Health Plans Study 2.0 survey. Med Care 2002;40:951-964.</p>
<p>27 (2d)</p>	<p>Measure Exclusions <i>Provide evidence to justify exclusion(s) and analysis of impact on measure results during testing.</i></p> <p>Summary of Evidence supporting exclusion(s):</p> <p>Citations for Evidence:</p> <p>Data/sample:</p> <p>Analytic Method:</p> <p>Testing Results:</p>
<p>28 (2e)</p>	<p>Risk Adjustment Testing <i>Summarize the testing used to determine the need (or no need) for risk adjustment and the statistical performance of the risk adjustment method.</i></p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis</p>

	<p>datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing abdominal aortic aneurysm repair.</p> <p>Analytic Method: Sensitivity analysis. We performed a sensitivity analysis to determine whether risk-adjustment of the mortality input was important in improving the predictive ability of the survival predictor measure. Risk-adjustment was performed using logistic regression to estimate expected mortality rates for each hospital based on patient age, gender, race, urgency of operation, median income, and coexisting diseases. Coexisting diseases were determined from secondary diagnostic codes using the methods of Elixhauser (16). The observed mortality rate at each hospital was then divided by the expected mortality rate to yield the ratio of observed/expected deaths (O/E ratio). The O/E ratio was multiplied by the average mortality rate for each operation to yield a risk-adjusted mortality rate. To determine the value of risk-adjustment in the context of selective referral, we compared the ability of risk-adjusted and unadjusted composite measures to predict subsequent performance.</p> <p>Testing Results: In sensitivity analysis, composite measures based on an unadjusted mortality input and a risk-adjusted mortality input had a correlation of (.95) and thus were equally good at predicting future performance (See pages 21-22 in the White Paper [3]).</p> <p>► If outcome or resource use measure not risk adjusted, provide rationale: Because risk-adjusted mortality is not available publicly except for limited locations, the capacity to use unadjusted mortality is very desirable, especially since it was shown to provide (under this methodology) an equal result. This measure will allow measurement to occur across the United States, providing information to national companies, health plans and consumers.</p>
<p>29 (2g)</p>	<p>Testing comparability of results when more than 1 data method is specified (e.g., administrative claims or chart abstraction) Data/sample: not applicable</p> <p>Analytic Method:</p> <p>Results:</p>
<p>30 (2f)</p>	<p>Provide Measure Results from Testing or Current Use Results from testing</p> <p>Data/sample: same as described above, results for survival predictor in White Paper [3] available on Website and Validation results for composite in [16] Staiger, Dimick et al., Medical Care 2009</p> <p>Methods to identify statistically significant and practically/meaningfully differences in performance: Bayesian Hierarchical methods using new shrinkage estimator Empirical Bayesian methods to determine weights Correlations Calculated the amount of variation predicted by survival predictor as a percentage of all hospital-level variation (adjusted for sampling variation)--analogous to a R-squared from a regression that summarizes the ability of the predictor to explain the hospital level variation in mortality for CABG surgery. Predictor was tested against the "gold standard" --risk adjusted mortality</p> <p>Results: See White Paper [3]</p>
<p>31 (2h)</p>	<p>Identification of Disparities ► If measure is stratified by factors related to disparities (i.e. race/ethnicity, primary language, gender, SES, health literacy), provide stratified results:</p> <p>► If disparities have been reported/identified, but measure is not specified to detect disparities, provide rationale: .</p>

USABILITY	
32 (3)	<p>Current Use Testing completed If in use, how widely used Nationally ▶ If "other," please describe: Survival Predictor for Pancreatectomy and Esophagectomy in use--see URL.</p> <p><input checked="" type="checkbox"/> Used in a public reporting initiative, name of initiative: Leapfrog Hospital Survey Sample report attached <input type="checkbox"/> OR Web page URL: https://www.leapfroggroup.org/cp</p>
33 (3a)	<p>Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)</p> <p>Data/sample:</p> <p>Methods:</p> <p>Results: See following citations reflecting consumer use of mortality information:</p> <p>[10]Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.</p> <p>[11]Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. Quality Review Bulletin 17:242-53.</p> <p>[12]Hibbard, J.H. and Jewett, J.(1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? Medical Care Research and Review., Vol 53(1): 28-47.</p>
34 (3b, 3c)	<p>Relation to other NQF-endorsed™ measures</p> <p>▶ Is this measure similar or related to measure(s) already endorsed by NQF (on the same topic or the same target population)? <i>Measures can be found at www.qualityforum.org under Core Documents.</i></p> <p><i>Check all that apply</i></p> <p><input type="checkbox"/> Have not looked at other NQF measures <input checked="" type="checkbox"/> Other measure(s) on same topic <input checked="" type="checkbox"/> Other measure(s) for same target population <input type="checkbox"/> No similar or related measures</p> <p>Name and number of similar or related NQF-endorsed™ measure(s): AHRQ AAA volume NQF#0357 --the Survival Predictor is focused at only elective AAA's--AHRQ's measure includes ruptured AAA's--the patient would not have an opportunity to select location for treatment with a ruptured aneurysm.</p> <p>AHRQ Risk-adjusted AVR Mortality NQF#0359--not aligned--the AHRQ measure includes emergent cases--the Survival Predictor is focused on non-ruptured cases where patients can determine where procedure will occur.</p> <p>Are the measure specifications harmonized with existing NQF-endorsed™ measures? Not harmonized</p> <p>▶ If not fully harmonized, provide rationale: This measure is focused at elective AAA repair..other measures are focused at elective and emergent. Given that we focused this measure at consumers and purchasers we wanted information about hospitals for selection of care provider.</p> <p>Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: This measure provides the ability to produce reliable mortality results for low volume hospitals, other measures do not have this capacity. In addition, the access to other mortality measures for AAA is focused on all-case mortality--this measure is focused solely on elective AAAs.</p>
FEASIBILITY	
35 (4a)	<p>How are the required data elements generated? <i>Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data elements are generated concurrent with and as a byproduct of care processes during care delivery (e.g., blood pressure or other assessment recorded by personnel conducting the assessment)</p> <p><input type="checkbox"/> Data elements are generated from a patient survey (e.g., CAHPS)</p> <p><input checked="" type="checkbox"/> Data elements are generated through coding performed by someone other than the person who obtained the original information (e.g., DRG or ICD-9 coding on claims)</p>

	<input checked="" type="checkbox"/> Other, Please describe: Data are currently submitted to Leapfrog via a secure online survey-
36 (4b)	<p>Electronic Sources All data elements</p> <p>► If all data elements are not in electronic sources, specify the near-term path to electronic collection by most providers:</p> <p>► Specify the data elements for the electronic health record: volume of AVR procedure, observed death during inpatient stay, following AVR procedure</p>
37 (4c)	<p>Do the specified exclusions require additional data sources beyond what is required for the other specifications? No</p> <p>► If yes, provide justification:</p>
38 (4d)	<p>Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure: It is unlikely that this procedure, or inpatient death will be inaccurately coded or not coded given the high cost of procedure and the accompanying death.</p> <p>Describe how could these potential problems be audited: If problems were identified, a chart review of cases could be performed.</p> <p>Did you audit for these potential problems during testing? No If yes, provide results:</p>
39 (4e)	<p>Testing feasibility Describe what have you learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: Initial results only available for Esophagectomy, Pancreatectomy. AAA, CABG, PCI, AVR will be released in 2009</p>
CONTACT INFORMATION	
40	<p>Web Page URL for Measure Information Describe where users (implementers) should go for more details on specifications of measures, or assistance in implementing the measure. Web page URL: https://leapfrog.medstat.com for access to Survival Predictor White Paper</p>
41	<p>Measure Steward Point of Contact</p> <p>First Name: MI: Last Name: Credentials (MD, MPH, etc.): Organization: The Leapfrog Group % The Academy Street Address: 1150 17th St., NW, Suite 600 City: Washington State: DC ZIP: 20036 Email: Telephone: ext:</p>
42	<p>Measure Developer Point of Contact If different from Measure Steward</p> <p>First Name: Justin MI: B Last Name: Dimick Credentials (MD, MPH, etc.): MD, MPH Organization: Department of Surgery, University of Michigan, M-SCORE offices, Suite 201 and 202 Street Address: 211 N. Fourth Avenue City: Ann Arbor State: MI ZIP: 48104 Email: jdimick@umich.edu Telephone: ext:</p>
ADDITIONAL INFORMATION	
43	<p>Workgroup/Expert Panel involved in measure development Workgroup/panel used</p> <p>► If workgroup used, describe the members' role in measure development: Research team led by Justin Dimick, MD, MPH;</p> <p>► Provide a list of workgroup/panel members' names and organizations:</p> <p>Douglas Staiger Ph.D., Department of Economics and the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth College, Hanover, New Hampshire</p> <p>John D. Birkmeyer, MD Michigan Surgical Collaborative for Outcomes Research and Evaluation</p>

	<p>Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Onur Baser, Ph.D. Michigan Surgical Collaborative for Outcomes Research and Evaluation Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Research supported by the National Institute on Aging</p>
44	<p>Measure Developer/Steward Updates and Ongoing Maintenance Year the measure was first released: 2008 Month and Year of most recent revision: August 2008 What is the frequency for review/update of this measure? Annual When is the next scheduled review/update for this measure? New coefficients for August 2009</p>
45	<p>Copyright statement/disclaimers: none</p>
46	<p>Additional Information: All measure information is available at https://leapfrog.medstat.com Please contact measure developer prior to use to assure all necessary items have been accessed.</p>
47	<p>I have checked that the submission is complete and any blank fields indicate that no information is provided. <input checked="" type="checkbox"/></p>
48	<p>Date of Submission (MM/DD/YY): Revised submission dated 3/18/09</p>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

The measure information you submit will be shared with NQF’s Steering Committees and Technical Advisory Panels to evaluate measures against the NQF criteria of importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Four conditions (as indicated below) must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, *importance to measure and report*, in order to be evaluated against the remaining criteria. References to the specific measure evaluation criteria are provided in parentheses following the item numbers. Please refer to the *Measure Evaluation Criteria* for more information at www.qualityforum.org under Core Documents. Additional guidance is being developed and when available will be posted on the NQF website.

Use the tab or arrow (↓→) keys to move the cursor to the next field (or back ←↑). There are three types of response fields:

- drop-down menus - select one response;
- check boxes - check as many as apply; and
- text fields - you can copy and paste text into these fields or enter text; these fields are not limited in size, but in most cases, we ask that you summarize the requested information.

Please note that URL hyperlinks do not work in the form; you will need to type them into your web browser.

Be sure to answer all questions. Fields that are left blank will be interpreted as no or none. **Information must be provided in this form.** Attachments are not allowed except to provide additional detail or source documents for information that is summarized in this form. If you have important information that is not addressed by the questions, they can be entered into item #46 near the end of the form.

For questions about this form, please contact the NQF Project Director listed in the corresponding call for measures.

CONDITIONS FOR CONSIDERATION BY NQF	
	Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards.
A (A)	Public domain or Measure Steward Agreement signed: Public domain - Agreement not required (If no, do not submit) <i>Template for the Measure Steward Agreement is available at www.qualityforum.org under Core Documents.</i>
B (B)	Measure steward/maintenance: Is there an identified responsible entity and process to maintain and update the measure on a schedule commensurate with clinical innovation, but at least every 3 years? Yes, information provided in contact section (If no, do not submit)
C (C)	Intended use: Does the intended use of the measure include BOTH public reporting AND quality improvement? Yes (If no, do not submit)
D (D)	Fully developed and tested: Is the measure fully developed AND tested? Yes, fully developed and tested (If not tested and no plans for testing within 24 months, do not submit)

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

	(for NQF staff use) NQF Review #: HOE-022-08 NQF Project: Hospital Outcomes and Efficiency
MEASURE SPECIFICATIONS & DESCRIPTIVE INFORMATION	
1	Information current as of (date- MM/DD/YY):
2	Title of Measure: Survival Predictor for Aortic Valve Replacement (AVR)
3	Brief description of measure ¹ : A reliability adjusted measure of AVR surgical performance that optimally combines two important domains: AVR hospital volume and AVR operative mortality, to provide predictions on AVR survival rates for hospitals. This measure is calculated based on data from administrative claims information.
4 (2a)	<p>Numerator Statement: Note: Because of the type of modeling done for this Survival Predictor--the information is not readily split into Numerator/ Denominator statements. Thus, we describe the two domains and their coding and data needs in this section.</p> <p>The formula for calculating the survival predictor has two components, one is a volume predicted mortality rate, and the second is an observed mortality rate. The volume predicted mortality rate reflects the hospitals experience performing AVR surgeries (thus, it includes all AVR surgeries) and uses mortality for all hospitals at that specific volume to create the volume predicted mortality. The input data from the hospitals for this domain is a volume count of all AVRs performed in the hospital.</p> <p>The second domain is the observed mortality, for this domain the population is the group of AVR cases, the data needed for this domain is the number of observed deaths occurring for AVR cases, within the inpatient setting.</p> <p>Note: All data is available in administrative claims information. In the case of Leapfrog's implementation hospitals are asked to submit aggregated information from their claims data. No personal health information is submitted to Leapfrog. Other users of the measure may have direct access to administrative data.</p> <p>Time Window: Annual</p> <p>Numerator Details (Definitions, codes with description): For the volume predicted mortality, hospitals count the number of AVR cases using the following codes:</p> <p>ICD-9-CM Procedure</p> <ul style="list-style-type: none"> ■ 35.21 Replacement of aortic valve with tissue graft ■ 35.22 Other replacement of aortic valve <p>See calculation worksheet for details on how volume-predicted mortality is used in the model.</p> <p>For the observed mortality domain, the hospital submits the observed deaths for AVR cases using the following codes (NQF endorsed)</p> <ul style="list-style-type: none"> ■ 35.21 Replacement of aortic valve with tissue graft ■ 35.22 Other replacement of aortic valve

¹ Example of measure description: Percentage of adult patients with diabetes aged 18-75 years receiving one or more A1c test(s) per year.
NQF Measure Submission Form, V3.1

	See Calculation Worksheet for examples of how the two domains are used to create the Survival Predictor.														
5 (2a)	<p>Denominator Statement:</p> <p>Time Window:</p> <p>Denominator Details (Definitions, codes with description):</p>														
6 (2a, 2d)	<p>Denominator Exclusions: No exclusions</p> <p>Denominator Exclusion Details (Definitions, codes with description): (</p>														
7 (2a, 2h)	<p>Stratification Do the measure specifications require the results to be stratified? No</p> <p>► If "other" describe:</p> <p>Identification of stratification variable(s):</p> <p>Stratification Details (Definitions, codes with description):</p>														
8 (2a, 2e)	<p>Risk Adjustment Does the measure require risk adjustment to account for differences in patient severity before the onset of care? No ► If yes, (select one)</p> <p>► Is there a separate proprietary owner of the risk model? No</p> <p>Identify Risk Adjustment Variables: See section 28 for rationale and support for not risk adjusting this measure. Measure was tested against risk adjusted mortality--details on that provided in Section 26.</p> <p>Detailed risk model: attached <input type="checkbox"/> OR Web page URL:</p>														
9 (2a)	<p>Type of Score: Rate/proportion Calculation Algorithm: attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p>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 = Score within a defined interval ► If "Other", please describe:</p>														
10 (2a, 4a, 4b)	<p>Identify the required data elements(e.g., primary diagnosis, lab values, vital signs): procedure codes</p> <p>Data dictionary/code table attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p>Data Quality (2a) Check all that apply</p> <p><input checked="" type="checkbox"/> Data are captured from an authoritative/accurate source (e.g., lab values from laboratory personnel)</p> <p><input checked="" type="checkbox"/> Data are coded using recognized data standards</p> <p><input checked="" type="checkbox"/> Method of capturing data electronically fits the workflow of the authoritative source</p> <p><input type="checkbox"/> Data are available in EHRs</p> <p><input checked="" type="checkbox"/> Data are auditable</p>														
11 (2a, 4b)	<p>Data Source and Data Collection Methods Identifies the data source(s) necessary to implement the measure specifications. Check all that apply</p> <table border="0"> <tr> <td><input type="checkbox"/> Electronic Health/Medical Record</td> <td><input type="checkbox"/> Paper Medical Record</td> </tr> <tr> <td><input type="checkbox"/> Electronic Clinical Database, Name:</td> <td><input type="checkbox"/> Standardized clinical instrument, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Clinical Registry, Name:</td> <td><input type="checkbox"/> Standardized patient survey, Name:</td> </tr> <tr> <td><input checked="" type="checkbox"/> Electronic Claims</td> <td><input type="checkbox"/> Standardized clinician survey, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Pharmacy data</td> <td><input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.</td> </tr> <tr> <td><input type="checkbox"/> Electronic Lab data</td> <td></td> </tr> <tr> <td><input type="checkbox"/> Electronic source - other, Describe:</td> <td></td> </tr> </table> <p>Instrument/survey attached <input checked="" type="checkbox"/> OR Web page URL:</p>	<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record	<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:	<input type="checkbox"/> Electronic Clinical Registry, Name:	<input type="checkbox"/> Standardized patient survey, Name:	<input checked="" type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:	<input type="checkbox"/> Electronic Pharmacy data	<input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.	<input type="checkbox"/> Electronic Lab data		<input type="checkbox"/> Electronic source - other, Describe:	
<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record														
<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:														
<input type="checkbox"/> Electronic Clinical Registry, Name:	<input type="checkbox"/> Standardized patient survey, Name:														
<input checked="" type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:														
<input type="checkbox"/> Electronic Pharmacy data	<input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.														
<input type="checkbox"/> Electronic Lab data															
<input type="checkbox"/> Electronic source - other, Describe:															
12 (2a)	<p>Sampling If measure is based on a sample, provide instructions and guidance on sample size.</p> <p>Minimum sample size:</p> <p>Instructions:</p>														

<p>13 (2a)</p>	<p>Type of Measure: Outcome ▶ If "Other", please describe:</p> <p>▶ If part of a composite or paired with another measure, please identify composite or paired measure While the measure uses two types of information components (domains), the results are not a composite as is defined by NQF, but rather a reliability adjusted measure of survival.</p> <p>Volume is used to create a volume predicted mortality for the hospital--this component of the measure is used to create greater reliability for low-volume hospitals. In the modeling for this measure, the volume predicted mortality and the observed mortality are weighted. In the model, lower volume hospitals have a higher weight on the volume predicted mortality versus the observed mortality. The opposite is true for high volume hospitals, which have a higher weight on the observed mortality. This methodology results in a reliability adjusted survival predictor.</p>
<p>14 (2a)</p>	<p>Unit of Measurement/Analysis (<i>Who or what is being measured</i>) Check all that apply.</p> <p><input type="checkbox"/> Can be measured at all levels <input type="checkbox"/> Individual clinician (e.g., physician, nurse) <input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice) <input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)</p> <p><input type="checkbox"/> Integrated delivery system <input checked="" type="checkbox"/> Health plan <input checked="" type="checkbox"/> Community/Population <input type="checkbox"/> Other (<i>Please describe</i>):</p>
<p>15 (2a)</p>	<p>Applicable Care Settings Check all that apply</p> <p><input type="checkbox"/> Can be used in all healthcare settings <input type="checkbox"/> Ambulatory Care (office/clinic) <input type="checkbox"/> Behavioral Healthcare <input type="checkbox"/> Community Healthcare <input type="checkbox"/> Dialysis Facility <input type="checkbox"/> Emergency Department <input type="checkbox"/> EMS emergency medical services <input type="checkbox"/> Health Plan <input type="checkbox"/> Home Health</p> <p><input type="checkbox"/> Hospice <input checked="" type="checkbox"/> Hospital <input type="checkbox"/> Long term acute care hospital <input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF) <input type="checkbox"/> Prescription Drug Plan <input type="checkbox"/> Rehabilitation Facility <input type="checkbox"/> Substance Use Treatment Program/Center <input type="checkbox"/> Other (<i>Please describe</i>):</p>
<p>IMPORTANCE TO MEASURE AND REPORT</p>	
<p>Note: This is a threshold criterion. If a measure is not judged to be sufficiently important to measure and report, it will not be evaluated against the remaining criteria.</p>	
<p>16 (1a)</p>	<p>Is measure related to a National Priority Partners priority area? Safety reliability (for NQF staff use) Does measure address a <u>specific</u> NPP goal? (www.qualityforum.org/about/NPP/):</p>
<p>17 (1a)</p>	<p>Does the measure address a high impact aspect of healthcare patient/societal consequences of poor quality</p> <p>Summary of Evidence: This measure addresses mortality in a high risk procedure (AVR) and is an outcome measure which is of interest to both consumers and purchasers.</p> <p>The rate of adjusted mortality for Medicare patients is relatively high for AVR compared to some surgeries ranging from 9.3% of cases for low volume hospitals to 7.1% of cases in high volume hospitals. [4b]</p> <p>In addition to addressing high volume procedure risk, this measure improves upon the technology of surgical procedure mortality measurement. It overcomes three problems with existing AVR mortality measures: 1) Mortality rates are often too "noisy" to reflect hospital quality with surgery (particularly among lower volume hospitals), 2) volume alone is a weak proxy for most procedures, and 3) when both volume and mortality are reported as separate indicators it is difficult to understand which measure is more important. [1] Given the large number of AVR procedures performed at low volume hospitals in the United States, and that this measure specifically addresses hospitals which perform elective procedures, consumers and purchasers would benefit from information that is more reliable in the prediction of future mortality for both selection and selective referral. In addition, this measure can be applied to the nation,</p>

	<p>states, or regions.</p> <p>Birkmeyer and Dimick (2009)[4] show that differences in mortality can be predicted using a reliability adjusted mortality rate (a weighted combination of volume and mortality) which is particularly relevant for selective-referral or public reporting contexts. They reduce the effects of random chance (statistical noise) and as a result with CABG, for example, more than half of the observed variation can be attributed to statistical noise. When they sorted hospitals simply on actual (risk-adjusted) mortality, rates varied from 1.4% to 11.0% across hospital quintiles (Figure 1 in White Paper [3]). After they adjusted for reliability, however, the mortality rates varied considerably less, from 3.3% to 6.3%.</p> <p>Citations² for Evidence:</p> <p>[1a] Rathore, A.J., Epstein, A.J., Rathore, S.S., Volpp, K.G., and Krunholz, H.M. (2004). Hospital percutaneous coronary intervention volume and patient mortality. 1988 to 2000: does the evidence support current procedure volume minimums? <i>J Am Coll Surg.</i>; 43:(10): 1755-62.</p> <p>[1b] Moscucci, M., Eagle, K.A., Share, D., Smith, D., DeFranco, A.C., O'Donnell, M., Kline-Rogers, E., Jani, S.M., and Brown, D.L. (2005). Public Reporting and Case Selection for Percutaneous Coronary Interventions: An Analysis from Two Large Multicenter Percutaneous Coronary Intervention Databases. <i>J. Am Coll Card.</i>, 45(11):1759-1765.</p> <p>[2] The National Hospital Bill: The Most Expensive Conditions by Payer, 2006. Statistical Brief #59. File accessed on March 16, 2009, at: http://www.hcup-us.ahrq.gov/reports/statbriefs/sb59.jsp Produced by AHRQ, Center for Delivery, Organization, and Markets, Healthcare Cost and Utilization Project, Nationwide Inpatient Sample, 2006.</p> <p>[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf</p> <p>[4] Birkmeyer, J.D., and Dimick, J.B. (2009) Understanding and reducing variation in surgical mortality. <i>Annu. Rev. Med.</i> 2009. 60:405-15.</p> <p>[4b] Birkmeyer, J.D., Siewers, A.E., Finlayson, E.V.A, Stukel, T.A., Lucas, F.L., Batista, I., Welch, G., Wennberg, D.A.. (2002) Hospital Volume and Surgical Mortality in US. <i>N Engl J Med</i>, Vol. 346, No. 15 •1128-1137.</p>
<p>18</p> <p>(1b)</p>	<p>Opportunity for Improvement <i>Provide evidence that demonstrates considerable variation, or overall poor performance, across providers.</i></p> <p>Summary of Evidence: In 2002, a systematic review of the literature on the volume-outcome relationship found that there was a significant relationship between hospital volume and outcomes. [5]</p> <p>The absolute differences in adjusted mortality rates between very-low volume hospitals and very-high-volume hospitals were slightly more than 2 percent for replacement of an aortic or mitral valve. Observed mortality varied across volume--at very low volume hospitals mortality rates were 9.9%, at very high volume hospitals observed mortality rates were 7.6%. Birkmeyer et al., 2002 [4b]</p> <p>Given the findings related to volume of procedures, Silber et al., [7] explored the relative contribution of complication rates and failure to rescue rates to mortality and found that complication rates were more likely influenced by patient factors while failure to rescue rates of those with complications was more related to hospital factors. Thus, it may be that higher volume hospitals are better at rescuing patients with complications. Silbers finding, in conjunction with the volume information, suggests lower volume hospitals with worse mortality rates could in fact address this through better care following the procedure, thereby reducing their overall rate. Unfortunately, most low volume hospitals in the United States do not have information on their AVRI mortality rate compared to other hospitals. When they are given this information, there is a good chance for improvement. Birkmeyer and Dimick [4] report that in</p>

² Citations can include, but are not limited to journal articles, reports, web pages (URLs).
NQF Measure Submission Form, V3.1

	<p>northern New England, mortality associated with CABG fell by >25% when hospitals and surgeons were given feedback on their mortality data. It is likely that feedback on AVR would have a similar impact.</p> <p>Note: Studies [4, 6] indicate it is also likely that some lower volume hospitals would also have lower mortality rates.</p> <p>Citations for Evidence:</p> <p>[1a] Rathore, A.J., Epstein, A.J., Rathore, S.S., Volpp, K.G., and Krumholz, H.M. (2004). Hospital percutaneous coronary intervention volume and patient mortality. 1988 to 2000: does the evidence support current procedure volume minimums? <i>J Am Coll Surg.</i>; 43:(10): 1755-62.</p> <p>[4b] Birkmeyer, J.D., Siewers, A.E., Finlayson, E.V.A., Stukel, T.A., Lucas, F.L., Batista, I., Welch, G., Wennberg, D.A.. (2002) Hospital Volume and Surgical Mortality in US. <i>N Engl J Med</i>, Vol. 346, No. 15 •1128-1137.</p> <p>[5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. <i>Annals of Internal Medicine</i>, Sept 1;137(6):511-20</p> <p>[6] Epstein, A.J., Rathore, S.S., Krumholz, H.M., and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>[7] Silber, J.H., Rosenbaum, P.R., Trudeau, M.E., et al. 2005. Changes in prognosis after the first postoperative complication. <i>Medical Care</i>, 43:122-31.</p>
<p>19 (1b)</p>	<p>Disparities <i>Provide evidence that demonstrates disparity in care/outcomes related to the measure focus among populations.</i></p> <p>Summary of Evidence: It is more likely that minorities will be treated at a low volume facility, and as a result are likely to be impacted by higher mortality rates. In an analysis of the National Inpatient Sample, Epstein, Rathore and Krumholz (2005)[6, pags 3-5] found that a greater proportion of patients treated in low volume hospitals for both CABG and PCI conditions were non-white, while a lower proportion of non-white patients presented as "elective" admissions or patients received in transfer as compared to patients in high volume hospitals. Given that about 10% of all CABG operations also include an AVR, it is likely that these findings for non-white patients would also hold for AVR.</p> <p>Citations for evidence:</p> <p>[6] Epstein, A.J., Rathore, S.S., Krumholz, H.M., and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p>
<p>20 (1c)</p>	<p>If measuring an Outcome Describe relevance to the national health goal/priority, condition, population, and/or care being addressed: <i>AVR is a high risk surgery, with mortality rates nearly 10% in some hospitals. [4b]</i></p> <p>This measure is highly relevant to both consumers and purchasers, given its frequency. National purchasers are interested in comparative information on hospitals nationwide. Pauly (1996) in a study of purchaser interests in hospital performance reporting found that mortality ratings were more important to purchasers than were morbidity or complications. [9] Health plans are interested in contracting with centers of excellence, which can be identified through the results of survival predictor in combination with other information on cost and quality. Consumers have shown their interest in cardiac procedure mortality by requesting reports from the state of Pennsylvania [10]; an earlier study by IOM (Lohr,</p>

Donaldson and Walker 1991) found that consumers were interested in hospital mortality rates, but did not perceive this information to be available.[11] Hibbard and Jewett found that consumers were more interested in "undesirable events" (such as mortality, complications, infections) than in "desirable events." [12]

[9] Pauly, M.V., Brailer, D.J.Kroch, E., and Even-Shoshan, O. Measuring Hospital Outcomes from a Buyer's Perspective. American Journal of Medical Quality, 11(8): Fall 1996.

[10] Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.

[11] Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. Quality Review Bulletin 17:242-53.

[12] Hibbard, J.H. and Jewett, J. (1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? Medical Care Research and Review., Vol 53(1): 28-47.

If not measuring an outcome, provide evidence supporting this measure topic and grade the strength of the evidence

Summarize the evidence (including citations to source) supporting the focus of the measure as follows:

- Intermediate outcome - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
- Process - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
- Structure - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
- Patient experience - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
- Access - evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
- Efficiency- demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

Type of Evidence Check all that apply

- | | |
|--|---|
| <input type="checkbox"/> Evidence-based guideline | <input checked="" type="checkbox"/> Quantitative research studies |
| <input checked="" type="checkbox"/> Meta-analysis | <input type="checkbox"/> Qualitative research studies |
| <input checked="" type="checkbox"/> Systematic synthesis of research | <input type="checkbox"/> Other (Please describe): |

Overall Grade for Strength of the Evidence³ (Use the USPSTF system, or if different, also describe how it relates to the USPSTF system): **Moderate**

Summary of Evidence (provide guideline information below): **Over 100 articles published related to volume and outcome relationship, with some inconsistency in results. Systematic review of the literature conducted in 2002. No review since that time.**

Citations for Evidence:

³The strength of the body of evidence for the specific measure focus should be systematically assessed and rated, e.g., USPSTF grading system www.ahrq.gov/clinic/uspstmeth.htm: **A** - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. **B** - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. **C** - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. **D** - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. **I** - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

	<p>[5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. <i>Annals of Internal Medicine</i>, Sept 1;137(6):511-20.</p> <p>[14] Birkmeyer, J.D., Dimick, J.B., Staiger, D.O. (2006) Operative mortality and procedure volume as predictors of subsequent hospital performance. <i>Ann Surg.</i> 243:411-417.</p> <p>[15] Dimick, JB, Welch HG, Birkmeyer JD. (2004) Surgical mortality as an indicator of hospital quality: The problem with small sample size. <i>JAMA</i>, 292:847-851.</p> <p>[4] Birkmeyer, JD., and Dimick, JB. (2009) Understanding and reducing variation in surgical mortality. <i>Annu. Rev. Med.</i> 60:405-15.</p> <p>[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. <i>Medical Care</i>, 47(2): 226-233.</p> <p>[6] Andrew J Epstein, Saif S Rathore, Harlan M Krumholz and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. <i>N Engl J Med.</i> 1979;301:1364-9.</p>
<p>21 (1c)</p>	<p>Clinical Practice Guideline <i>Cite the guideline reference; quote the specific guideline recommendation related to the measure and the guideline author's assessment of the strength of the evidence; and summarize the rationale for using this guideline over others.</i></p> <p>Guideline Citation:</p> <p>Specific guideline recommendation:</p> <p>Guideline author's rating of strength of evidence (If different from USPSTF, also describe it and how it relates to USPSTF):</p> <p>Rationale for using this guideline over others: .</p>
<p>22 (1c)</p>	<p>Controversy/Contradictory Evidence <i>Summarize any areas of controversy, contradictory evidence, or contradictory guidelines and provide citations.</i></p> <p>Summary: There are three areas of possible contention with the survival predictor measure--</p> <p>1) The volume-outcome relationship has been questioned for some procedures [6, 17, 19, 25]</p> <p>Epstein and Rathore et al., [6] questioned whether it was appropriate to move patients from low volume hospitals to high volume hospitals, given the number of patients that would have to be moved to save 1 life. They did find in their study of CABG and PCI's performed in the US that low volume hospitals did have higher unadjusted and adjusted for case mix mortality. Of concern, is that 38% of all CABG surgery is performed in low volume hospitals; and that non-white patients were more likely to be treated at low volume hospitals.</p> <p>Peterson et al., [19] questioned the volume outcomes relationship for CABG surgery, and found only modest associations for volume and outcome for CABG. Those with high volume had mortality rate of 2.5% while low volume hospitals rate was 3.2%. They suggest using past mortality rate to select hospitals. (The survival predictor uses both volume and mortality to predict survival in the next year.)</p> <p>Yet, more than 100 studies have demonstrated better results at high-volume hospitals with cardiovascular surgery, major cancer resections, and other high-risk procedures.[18, 20]</p> <p>There is specific evidence of the variation in mortality across the different volume levels of AVRs. [4b]. They documented that there were differences between low volume hospitals and high volume hospitals in</p>

	<p>mortality--with high volume hospitals having less mortality.</p> <p>2) That outcome measures must be risk-adjusted unless there is evidence to show it is not needed (NQF). The survival predictor measure predicts better than volume or mortality alone, and is as good a predictor as risk-adjusted mortality. When testing the unadjusted survival predictor against risk-adjusted mortality there was a (.96) correlation. [4] See Section 28 of this form for details.</p> <p>3) The weighting of input measures into composites. Existing approaches rely on overly simplistic approaches. Among these, assigning equal weight to all measures (i.e., the all or none approach) and relying on expert opinion are the most common. The survival predictor relies on empiric methods for weighting the input measures.</p> <p>Citations:</p> <p>[1a] Rathore, A.J., Epstein, A.J., Rathore, S.S., Volpp, K.G., and Krunholz, H.M. (2004). Hospital percutaneous coronary intervention volume and patient mortality. 1988 to 2000: does the evidence support current procedure volume minimums? <i>J Am Coll Surg.</i>; 43:(10): 1755-62.</p> <p>[6] Andrew J Epstein, Saif S Rathore, Harlan M Krumholz and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>[16]Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. <i>Medical Care</i>, 47(2): p. 232.</p> <p>[17] Edward L. Hannan, PhD; Chuntao Wu, PhD; Thomas J. Ryan, MD; Edward Bennett, MD; Alfred T. Culliford, MD; Jeffrey P. Gold, MD; Alan Hartman, MD; O. Wayne Isom, MD; Robert H. Jones, MD; Barbara McNeil, MD, PhD; Eric A. Rose, MD; Valavanur A. Subramanian, MD. Do Hospitals and Surgeons With Higher Coronary Artery Bypass Graft Surgery Volumes Still Have Lower Risk-Adjusted Mortality Rates? <i>Circulation</i>. 2003;108:795-801.</p> <p>[18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. <i>N Engl J Med</i>. 1979;301:1364-9.</p> <p>[19] Eric D. Peterson, MD, MPH; Laura P. Coombs, PhD; Elizabeth R. DeLong, PhD; Constance K. Haan, MD; T. Bruce Ferguson, MD. Procedural Volume as a Marker of Quality for CABG Surgery. <i>JAMA</i>. 2004;291:195-201.</p> <p>[20] Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. <i>JAMA</i>. 1998;280:1747-51.)</p> <p>[25] McGrath, PD., Wennberg, DE., Dickens, Jr., JD., Siewers, AE., Lucas, FL., Malenka, DJ, Kellett, Jr., MA., Ryan, Jr., TJ. (2000) Relation Between Operator and Hospital Volume and Outcomes Following Percutaneous Coronary Interventions in the Era of the Coronary Stent. <i>JAMA</i>, 284(24):3139-3144.</p>
<p>23 (1)</p>	<p>Briefly describe how this measure (as specified) will facilitate significant gains in healthcare quality related to the specific priority goals and quality problems identified above:</p> <p>This measure of predicted survival improves upon the reliability of mortality results for high risk surgical procedures, such as AVR. For the first time, this measure produces reliable mortality/survivability information on smaller volume hospitals, as well as high volume hospitals. Hospitals across the country will have information available through voluntary public reporting.</p>
<p>SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES</p>	
<p>Note: Testing and results should be summarized in this form. However, additional detail and reports may be submitted as supplemental information or provided as a web page URL. If a measure has not been tested, it is only potentially eligible for time-limited endorsement.</p>	

24	Supplemental Testing Information: attached <input checked="" type="checkbox"/> OR Web page URL:
25	<p>Reliability Testing</p> <p>(2b) Data/sample: Data was a 100% sample from the Medicare Analysis Provider and Review (MEDPAR) files for 2000-2003, these files contain 100% of Medicare hospitalizations for years specified. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing an aortic valve replacement. Codes used to select patients are indicated in Section 4 of this form.</p> <p>This dataset has been used by other researchers to look at the issue of volume and mortality for PCI. As indicated by McGrath and Wennberg et al., [25] the Medicare dataset allowed sufficient power to determine significant differences in adverse outcomes across varying levels of volume.</p> <p>Note: Needleman, Buerhaus, et al., (2003) concluded after applying operational tests on Medicare data for adverse outcomes and all-patient hospital data from 11 states, that Medicare data could be used to assess quality in hospitals.[20] Given the lack of a national all-patient/all-payer database, MEDPAR data was used in development and testing of the models.</p> <p>Analytic Method:</p> <p>Model Development</p> <p>We used an empirical Bayes approach to combine mortality rates with information on hospital volume at each hospital. In traditional empirical Bayes methods, a point estimate (e.g., mortality rate observed at a hospital) is adjusted for reliability by shrinking it towards the overall mean (e.g., overall mortality rate in the population) [21,22]. We modified this traditional approach by shrinking the observed mortality rate back toward the mortality rate expected given the volume at that hospital—we refer to this as the “volume-predicted mortality” (See attached White Paper [3] TECHNICAL APPENDIX for the mathematical details of this method). With this approach, the observed mortality rate is weighted according to how reliably it is estimated, with the remaining weight placed on the information regarding hospital volume. Because this method includes observed data to the extent that it is useful, and only relies on the proxy measure to the extent necessary, it ensures an optimal combination of these two quality domains. [3]</p> <p>The two inputs to the survival predictor measure are mortality rates and procedure volume for each of the six included operations. Procedure-specific mortality rates were calculated for all hospitals over a 2-year period (2000-01) and this was used as the first input. Hospital volume was calculated as the number of Medicare cases performed during the same time period. For each operation, the relationship between hospital volume and risk-adjusted mortality was modeled using linear regression. (Details of the risk-adjustment strategy will be discussed below.) After testing the fit of several transformations, hospital volume was modeled as the natural log of the continuous volume variable, which is the same approach used in our previous work [23]. Using this regression model, we estimated the volume-predicted mortality, the second input to the survival predictor measure.</p> <p>We then used the empirical Bayes approach to create an optimal combination of these two inputs. This survival predictor measure theoretically provides the best estimate of a hospital's true mortality rate, taking into account the both available inputs [21,22]. The combined survival predictor measure was calculated as follows: mortality prediction = (weight)*(observed mortality) + (1-weight)*(volume-predicted mortality). The weight placed on the point estimate of mortality is the reliability, or ratio of signal to signal plus noise, calculated as follows: weight = variation among hospitals/(variation among hospitals + variation within hospitals). The variation among hospitals was calculated as the variance in observed</p>

	<p>mortality rates for the hospitals included in the sample. The variation within hospitals was calculated as the standard error of the mortality rate at each hospital. With this method, more weight is placed on the observed mortality rate when a hospital has a high number of cases because it is estimated with more reliability; less weight is placed on the observed mortality rate when a hospital performs a low number of cases because of its lower reliability.</p> <p>A calculation worksheet with examples is attached.</p> <p>Testing Results: Hospital caseloads and the weights applied to each input to the survival predictor measure varied for each procedure studied (see Table 1 in white paper [3]). For aortic valve replacement, a procedure with lower hospital caseloads than CABG or PCI, the weight applied to the volume input was .73. ([3]-Table 1). For hospitals with higher volumes more weight was placed on the observed mortality.</p> <p>The survival predictor (mortality) measure explained a large proportion of non-random, hospital-level variation in risk-adjusted mortality rates (see Table 2, p. 19 in White Paper [3]). For aortic valve replacement the survival predictor explained 47% of the hospital level variation in mortality rates; this compares to 26% for observed mortality and 18% for volume of AVR cases. Measures with low reliability or correlation explain little variation. The correlation between the survival predictor and risk-adjusted mortality was (.96) ([16] p. 232), and the amount of variation explained for AVR was 47% [3]. This is a more than adequate level of reliability.</p> <p>Note: The percentage of hospital level variation in mortality rates explained by the survival predictor is analogous to R squared in regression analysis. [16, p. 228]</p> <p>Citations:</p> <p>[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf</p> <p>[21] Morris CN. Parametric Empirical Bayes Inference: Theory and Applications. J Am Stat Assoc 1988;78:47-55.</p> <p>[22] McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) Frontiers in Health Policy Research. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.</p> <p>[23] Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. N Engl J Med. 2003;349:2117-2127.</p> <p>[20] Needleman, J., Buerhaus, P.I., Mattke, S., Stewart, M., and Zelevinsky, M. (2003). Health Services Research 38.6, Part I; 1487-1508.</p> <p>[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. Medical Care, 47(2): 226-233.</p> <p>[25] McGrath, PD., Wennberg, DE., Dickens, Jr., JD., Siewers, AE., Lucas, FL., Malenka, DJ, Kellett, Jr., MA., Ryan, Jr., TJ. (2000) Relation Between Operator and Hospital Volume and Outcomes Following Percutaneous Coronary Interventions in the Era of the Coronary Stent. JAMA, 284(24):3139-3144.</p>
<p>26</p> <p>(2c)</p>	<p>Validity Testing</p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p>

Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the aortic valve replacement surgery.

Analytic Method: We determined the value of our survival predictor (mortality) measure by establishing whether it explained hospital-level variation in risk-adjusted mortality rates and by assessing to what degree it was able to predict future hospital performance. We first estimated the proportion of variation in hospital-level mortality (2000-01) explained by the survival predictor measure using random effects logistic regression models. For these analyses, we estimated the proportional change in the hospital-level variance in mortality rates, which was determined from the standard deviation of the random effect, after adding each measure to the model [14,22].

We next compared the ability of the survival predictor measure to the individual measures, mortality rates and hospital volume. We should note that these analyses focus on explaining systematic, or non-random, variation, since measurement error (random error) is accounted for and subtracted from the total variation in all analyses [22,24].

We next determined the extent to which the composite measure predicts future risk-adjusted mortality. For this analysis, hospitals were ranked based on each measure from the earlier time period (data from years 2000-01) and divided into four equal size groups (quartiles at the patient level). The subsequent risk-adjusted mortality rates for each quartile of performance were then calculated (data from years 2002-03). We present the subsequent mortality rates across quartiles of the AVR survival predictor measure to graphically demonstrate its usefulness in discriminating among hospitals for the entire spectrum of performance. To compare the predictive ability of the composite measures and individual measures, we also present the subsequent mortality rates in the "worst" compared to the "best" quartile in the White Paper ([3], p. 22) "Quartiles of Performance Measures (2000-2001). This table reflects how well the unadjusted survival predictor created on 2000-2001 data compares to risk-adjusted mortality in 2002-2003 data.

Note: The risk-adjusted mortality rate for AVR was constructed using standard methods. We determined the ratio of actual deaths or complications to the number of expected deaths (the O/E ratio). The number of expected deaths was the sum over all patients of the predicted probability of death or complications derived from a logistic regression model estimated on all patients undergoing PCI. The dependent variable in the logistic model was death or complications and the independent variables were patient covariates. The patient characteristics included age, gender, race, admission acuity, and co-existing diseases using the Elixhauser method. A zip code level measure of socio-economic status was derived from 2000 census data.

Testing Results: While some measures are good at discriminating top performers or bottom performers, this measure is good at prediction across entire spectrum of performance. [See White paper [3]: Figures p. 21-22) for a graphical demonstration of the usefulness of the survival predictor in discriminating among hospitals across the entire spectrum of performance.] To compare the predictive ability of the reliability adjusted survival predictor versus the individual components (volume and observed mortality) we also present the subsequent mortality rates in the "worst" compared to the "best" quartile. In the case of AVR, the Survival Predictor was a better predictor of subsequent risk adjusted mortality than either hospital volume alone or observed mortality alone. [3, p. 20]

[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: <http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf>

[22]. McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) Frontiers in Health Policy Research. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.

[14] Birkmeyer JD, Dimick JB, Staiger DO. Operative mortality and procedure volume as predictors of subsequent hospital performance. Ann Surg 2006;243:411-417.

	[24] Zaslavsky AM, Cleary PD. Dimensions of plan performance for sick and healthy members on the Consumer Assessments of Health Plans Study 2.0 survey. <i>Med Care</i> 2002;40:951-964.
27 (2d)	<p>Measure Exclusions <i>Provide evidence to justify exclusion(s) and analysis of impact on measure results during testing.</i></p> <p>Summary of Evidence supporting exclusion(s):</p> <p>Citations for Evidence:</p> <p>Data/sample:</p> <p>Analytic Method:</p> <p>Testing Results:</p>
28 (2e)	<p>Risk Adjustment Testing <i>Summarize the testing used to determine the need (or no need) for risk adjustment and the statistical performance of the risk adjustment method.</i></p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing aortic valve replacement.</p> <p>Analytic Method: Sensitivity analysis. We performed a sensitivity analysis to determine whether risk-adjustment of the mortality input was important in improving the predictive ability of the survival predictor measure. Risk-adjustment was performed using logistic regression to estimate expected mortality rates for each hospital based on patient age, gender, race, urgency of operation, median income, and coexisting diseases. Coexisting diseases were determined from secondary diagnostic codes using the methods of Elixhauser (16). The observed mortality rate at each hospital was then divided by the expected mortality rate to yield the ratio of observed/expected deaths (O/E ratio). The O/E ratio was multiplied by the average mortality rate for each operation to yield a risk-adjusted mortality rate. To determine the value of risk-adjustment in the context of selective referral, we compared the ability of risk-adjusted and unadjusted composite measures to predict subsequent performance.</p> <p>Testing Results: In sensitivity analysis, composite measures based on an unadjusted mortality input and a risk-adjusted mortality input had a correlation of (.95) and thus were equally good at predicting future performance (See pages 21-22 in the White Paper [3]).</p> <p>► If outcome or resource use measure not risk adjusted, provide rationale: Because risk-adjusted AVR mortality is not available publicly except for limited locations, the capacity to use unadjusted mortality is very desirable, especially since it was shown to provide (under this methodology) an equal result. This measure will allow measurement to occur across the United States, providing information to national companies, health plans and consumers.</p>
29 (2g)	<p>Testing comparability of results when more than 1 data method is specified (<i>e.g., administrative claims or chart abstraction</i>)</p> <p>Data/sample: not applicable</p> <p>Analytic Method:</p> <p>Results:</p>
30	Provide Measure Results from Testing or Current Use Results from testing

(2f)	<p>Data/sample: same as described above, results for survival predictor in White Paper [3]available on Website and Validation results for composite in [16] Staiger, Dimick et al., Medical Care 2009</p> <p>Methods to identify statistically significant and practically/meaningfully differences in performance: Bayesian Hierarchical methods using new shrinkage estimator Empirical Bayesian methods to determine weights Correlations Calculated the amount of variation predicted by survival predictor as a percentage of all hospital-level variation (adjusted for sampling variation)--analogous to a R-squared from a regression that summarizes the ability of the predictor to explain the hospital level variation in mortality for AVR surgery. Predictor was tested against the "gold standard" --risk adjusted mortality</p> <p>Results: See White Paper [3]</p>
31 (2h)	<p>Identification of Disparities ► If measure is stratified by factors related to disparities (i.e. race/ethnicity, primary language, gender, SES, health literacy), provide stratified results:</p> <p>► If disparities have been reported/identified, but measure is not specified to detect disparities, provide rationale: .</p>
USABILITY	
32 (3)	<p>Current Use Testing completed If in use, how widely used Nationally ► If "other," please describe: Survival Predictor for Pancreatectomy and Esophagectomy in use--see URL.</p> <p><input checked="" type="checkbox"/> Used in a public reporting initiative, name of initiative: Leapfrog Hospital Survey Sample report attached <input type="checkbox"/> OR Web page URL: https://www.leapfroggroup.org/cp</p>
33 (3a)	<p>Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)</p> <p>Data/sample:</p> <p>Methods:</p> <p>Results: See following citations reflecting consumer use of mortality information:</p> <p>[10]Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.</p> <p>[11]Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. Quality Review Bulletin 17:242-53.</p> <p>[12]Hibbard, J.H. and Jewett, J.(1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? Medical Care Research and Review., Vol 53(1): 28-47.</p>
34 (3b, 3c)	<p>Relation to other NQF-endorsed™ measures ► Is this measure similar or related to measure(s) already endorsed by NQF (on the same topic or the same target population)? <i>Measures can be found at www.qualityforum.org under Core Documents.</i> Check all that apply</p> <p><input type="checkbox"/> Have not looked at other NQF measures <input checked="" type="checkbox"/> Other measure(s) on same topic <input checked="" type="checkbox"/> Other measure(s) for same target population <input type="checkbox"/> No similar or related measures</p> <p>Name and number of similar or related NQF-endorsed™ measure(s): CMS AVR volume NQF#0124--Survival Predictor for AVR aligned with CMS measure specifications Risk-adjusted AVR Mortality NQF#0120--not aligned--risk adjustment methodology requires registry data--no access to data available.</p>

	<p>Are the measure specifications harmonized with existing NQF-endorsed™ measures? <i>Partially harmonized</i></p> <p>► If not fully harmonized, provide rationale: <i>This new measure requires combination of volume and mortality--no other measure uses this combination. The other endorsed mortality measure requires registry information to complete the risk adjustment; we are harmonized with specifications for volume with CMS endorsed measure.</i></p> <p>Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: <i>This measure provides the ability to produce reliable mortality results for low volume hospitals, other measures do not have this capacity. In addition, the access to data nationally for other AVR mortality measures does not exist.</i></p>
FEASIBILITY	
35 (4a)	<p>How are the required data elements generated? <i>Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data elements are generated concurrent with and as a byproduct of care processes during care delivery (e.g., blood pressure or other assessment recorded by personnel conducting the assessment)</p> <p><input type="checkbox"/> Data elements are generated from a patient survey (e.g., CAHPS)</p> <p><input checked="" type="checkbox"/> Data elements are generated through coding performed by someone other than the person who obtained the original information (e.g., DRG or ICD-9 coding on claims)</p> <p><input checked="" type="checkbox"/> Other, Please describe: <i>Data are currently submitted to Leapfrog via a secure online survey-</i></p>
36 (4b)	<p>Electronic Sources <i>All data elements</i></p> <p>► If all data elements are not in electronic sources, specify the near-term path to electronic collection by most providers:</p> <p>► Specify the data elements for the electronic health record: <i>volume of AVR procedure, observed death during inpatient stay, following AVR procedure</i></p>
37 (4c)	<p>Do the specified exclusions require additional data sources beyond what is required for the other specifications? <i>No</i></p> <p>► If yes, provide justification:</p>
38 (4d)	<p>Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure: <i>It is unlikely that this procedure, or inpatient death will be inaccurately coded or not coded given the high cost of procedure and the accompanying death.</i></p> <p>Describe how could these potential problems be audited: <i>If problems were identified, a chart review of cases could be performed.</i></p> <p>Did you audit for these potential problems during testing? <i>No</i> If yes, provide results:</p>
39 (4e)	<p>Testing feasibility Describe what have you learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: <i>Initial results only available for Esophagectomy, Pancreatectomy. AVR, AAA, PCI and CABG will be released in 2009</i></p>
CONTACT INFORMATION	
40	<p>Web Page URL for Measure Information <i>Describe where users (implementers) should go for more details on specifications of measures, or assistance in implementing the measure.</i></p> <p>Web page URL: https://leapfrog.medstat.com for access to Survival Predictor White Paper</p>
41	<p>Measure Steward Point of Contact</p> <p>First Name: MI: Last Name: Credentials (MD, MPH, etc.):</p> <p>Organization: <i>The Leapfrog Group % The Academy</i></p> <p>Street Address: <i>1150 17th St., NW, Suite 600</i> City: <i>Washington</i> State: <i>DC</i> ZIP: <i>20036</i></p>

	Email: Telephone: ext:
42	<p>Measure Developer Point of Contact If different from Measure Steward</p> <p>First Name: Justin MI: B Last Name: Dimick Credentials (MD, MPH, etc.): MD, MPH</p> <p>Organization: Department of Surgery, University of Michigan, M-SCORE offices, Suite 201 and 202</p> <p>Street Address: 211 N. Fourth Avenue City: Ann Arbor State: MI ZIP: 48104</p> <p>Email: jdimick@umich.edu Telephone: ext:</p>
ADDITIONAL INFORMATION	
43	<p>Workgroup/Expert Panel involved in measure development Workgroup/panel used</p> <p>▶ If workgroup used, describe the members' role in measure development: Research team led by Justin Dimick, MD, MPH;</p> <p>▶ Provide a list of workgroup/panel members' names and organizations:</p> <p>Douglas Staiger Ph.D., Department of Economics and the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth College, Hanover, New Hampshire</p> <p>John D. Birkmeyer, MD Michigan Surgical Collaborative for Outcomes Research and Evaluation Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Onur Baser, Ph.D. Michigan Surgical Collaborative for Outcomes Research and Evaluation Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Research supported by the National Institute on Aging</p>
44	<p>Measure Developer/Steward Updates and Ongoing Maintenance</p> <p>Year the measure was first released: 2008</p> <p>Month and Year of most recent revision: August 2008</p> <p>What is the frequency for review/update of this measure? Annual</p> <p>When is the next scheduled review/update for this measure? New coefficients for August 2009</p>
45	Copyright statement/disclaimers: none
46	Additional Information: All measure information is available at https://leapfrog.medstat.com Please contact measure developer prior to use to assure all necessary items have been accessed.
47	I have checked that the submission is complete and any blank fields indicate that no information is provided. <input checked="" type="checkbox"/>
48	Date of Submission (MM/DD/YY): Revised submission dated 3/18/09

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

The measure information you submit will be shared with NQF’s Steering Committees and Technical Advisory Panels to evaluate measures against the NQF criteria of importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Four conditions (as indicated below) must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, *importance to measure and report*, in order to be evaluated against the remaining criteria. References to the specific measure evaluation criteria are provided in parentheses following the item numbers. Please refer to the *Measure Evaluation Criteria* for more information at www.qualityforum.org under Core Documents. Additional guidance is being developed and when available will be posted on the NQF website.

Use the tab or arrow (↓→) keys to move the cursor to the next field (or back ←↑). There are three types of response fields:

- drop-down menus - select one response;
- check boxes - check as many as apply; and
- text fields - you can copy and paste text into these fields or enter text; these fields are not limited in size, but in most cases, we ask that you summarize the requested information.

Please note that URL hyperlinks do not work in the form; you will need to type them into your web browser.

Be sure to answer all questions. Fields that are left blank will be interpreted as no or none. **Information must be provided in this form.** Attachments are not allowed except to provide additional detail or source documents for information that is summarized in this form. If you have important information that is not addressed by the questions, they can be entered into item #46 near the end of the form.

For questions about this form, please contact the NQF Project Director listed in the corresponding call for measures.

CONDITIONS FOR CONSIDERATION BY NQF	
	Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards.
A (A)	Public domain or Measure Steward Agreement signed: Public domain - Agreement not required (If no, do not submit) <i>Template for the Measure Steward Agreement is available at www.qualityforum.org under Core Documents.</i>
B (B)	Measure steward/maintenance: Is there an identified responsible entity and process to maintain and update the measure on a schedule commensurate with clinical innovation, but at least every 3 years? Yes, information provided in contact section (If no, do not submit)
C (C)	Intended use: Does the intended use of the measure include BOTH public reporting AND quality improvement? Yes (If no, do not submit)
D (D)	Fully developed and tested: Is the measure fully developed AND tested? Yes, fully developed and tested (If not tested and no plans for testing within 24 months, do not submit)

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

	(for NQF staff use) NQF Review #: HOE-023-08 NQF Project: Hospital Outcomes and Efficiency
MEASURE SPECIFICATIONS & DESCRIPTIVE INFORMATION	
1	Information current as of (date- MM/DD/YY):
2	Title of Measure: Survival Predictor for Esophagectomy Surgery
3	Brief description of measure ¹ : A reliability adjusted measure of Esophagectomy surgical performance that optimally combines two important domains: Esophagectomy hospital volume and Esophagectomy operative mortality, to provide predictions on Esophagectomy survival rates for hospitals. This measure is calculated based on data from administrative claims information.
4 (2a)	<p>Numerator Statement: Note: Because of the type of modeling done for this Survival Predictor--the information is not readily split into Numerator/ Denominator statements. Thus, we describe the two domains and their coding and data needs in this section.</p> <p>The formula for calculating the survival predictor has two components, one is a volume predicted mortality rate, and the second is an observed mortality rate. The volume predicted mortality rate reflects the hospitals experience performing Esophagectomy surgeries (thus, it includes all Esophagectomy surgeries) and uses mortality for all hospitals at that specific volume to create the volume predicted mortality. The input data from the hospitals for this domain is a volume count of all Esophagectomys performed in the hospital.</p> <p>The second domain is the observed mortality, for this domain the population is narrowed to a homogenous group of esophagectomy with a diagnosis of cancer, the data needed for this domain is the number of observed deaths occurring for esophagectomy cases with cancer, within the inpatient setting.</p> <p>Note: All data is available in administrative claims information. In the case of Leapfrog's implementation hospitals are asked to submit aggregated information from their claims data. No personal health information is submitted to Leapfrog. Other users of the measure may have direct access to administrative data.</p> <p>Time Window: 12 months</p> <p>Numerator Details (Definitions, codes with description): For the volume predicted mortality, hospitals count the number of esophagectomy cases using the following codes:</p> <p>ICD-9-CM Procedure Codes:</p> <p>424 Esophagectomy 4240 Esophagectomy NOS 4241 Partial Esophagectomy 4242 Total Esophagectomy 4399 Total gastrectomy NEC</p> <p>See calculation worksheet for details on how volume-predicted mortality is used in the model.</p> <p>For the observed mortality domain, the hospital submits the observed deaths for esophagectomy cases</p>

¹ Example of measure description: Percentage of adult patients with diabetes aged 18-75 years receiving one or more A1c test(s) per year.
NQF Measure Submission Form, V3.1

	<p>with a cancer diagnosis using the following codes:</p> <p>ICD-9-CM Procedure Codes:</p> <p>424 Esophagectomy 4240 Esophagectomy NOS 4241 Partial Esophagectomy 4242 Total Esophagectomy 4399 Total gastrectomy NEC</p> <p>And, one of the following esophageal cancer diagnoses:</p> <p>1500 MAL NEO CERVICAL ESOPHAG 1501 MAL NEO THORACIC ESOPHAG 1502 MAL NEO ABDOMIN ESOPHAG 1503 MAL NEO UPPER 3RD ESOPH 1504 MAL NEO MIDDLE 3RD ESOPH 1505 MAL NEO LOWER 3RD ESOPH 1508 MAL NEO ESOPHAGUS NEC 1509 MAL NEO ESOPHAGUS NOS</p> <p>Thus, the observed mortality is based on the volume count of esophagectomys and an actual count of deaths occurring for that subset of esophagectomys with cancer as a diagnosis.</p> <p>See Calculation Worksheet for how the two domains are used to create the Survival Predictor.</p>
5 (2a)	<p>Denominator Statement: See numerator section for all data needed, and codes</p> <p>Time Window:</p> <p>Denominator Details (Definitions, codes with description):</p>
6 (2a, 2d)	<p>Denominator Exclusions: None</p> <p>Denominator Exclusion Details (Definitions, codes with description):</p>
7 (2a, 2h)	<p>Stratification Do the measure specifications require the results to be stratified? No ► If "other" describe:</p> <p>Identification of stratification variable(s):</p> <p>Stratification Details (Definitions, codes with description):</p>
8 (2a, 2e)	<p>Risk Adjustment Does the measure require risk adjustment to account for differences in patient severity before the onset of care? No ► If yes, (select one) ► Is there a separate proprietary owner of the risk model? No</p> <p>Identify Risk Adjustment Variables: See section 28 for rationale and support for not risk adjusting this measure. Measure was tested against risk adjusted mortality--details on that provided in Section 26.</p> <p>Detailed risk model: attached <input type="checkbox"/> OR Web page URL:</p>
9 (2a)	<p>Type of Score: Rate/proportion Calculation Algorithm: attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p>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 = Score within a defined interval ► If "Other", please describe:</p>
10	<p>Identify the required data elements(e.g., primary diagnosis, lab values, vital signs): procedure codes, diagnosis codes</p>

(2a, 4a, 4b)	<p>Data dictionary/code table attached <input checked="" type="checkbox"/> OR Web page URL:</p> <p>Data Quality (2a) <i>Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data are captured from an authoritative/accurate source (e.g., lab values from laboratory personnel)</p> <p><input checked="" type="checkbox"/> Data are coded using recognized data standards</p> <p><input checked="" type="checkbox"/> Method of capturing data electronically fits the workflow of the authoritative source</p> <p><input type="checkbox"/> Data are available in EHRs</p> <p><input checked="" type="checkbox"/> Data are auditable</p>																		
11 (2a, 4b)	<p>Data Source and Data Collection Methods <i>Identifies the data source(s) necessary to implement the measure specifications. Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Electronic Health/Medical Record</td> <td><input type="checkbox"/> Paper Medical Record</td> </tr> <tr> <td><input type="checkbox"/> Electronic Clinical Database, Name:</td> <td><input type="checkbox"/> Standardized clinical instrument, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Clinical Registry, Name:</td> <td><input type="checkbox"/> Standardized patient survey, Name:</td> </tr> <tr> <td><input checked="" type="checkbox"/> Electronic Claims</td> <td><input type="checkbox"/> Standardized clinician survey, Name:</td> </tr> <tr> <td><input type="checkbox"/> Electronic Pharmacy data</td> <td><input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.</td> </tr> <tr> <td><input type="checkbox"/> Electronic Lab data</td> <td></td> </tr> <tr> <td><input type="checkbox"/> Electronic source - other, Describe:</td> <td></td> </tr> </table> <p style="text-align: right;">Instrument/survey attached <input checked="" type="checkbox"/> OR Web page URL:</p>	<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record	<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:	<input type="checkbox"/> Electronic Clinical Registry, Name:	<input type="checkbox"/> Standardized patient survey, Name:	<input checked="" type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:	<input type="checkbox"/> Electronic Pharmacy data	<input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.	<input type="checkbox"/> Electronic Lab data		<input type="checkbox"/> Electronic source - other, Describe:					
<input type="checkbox"/> Electronic Health/Medical Record	<input type="checkbox"/> Paper Medical Record																		
<input type="checkbox"/> Electronic Clinical Database, Name:	<input type="checkbox"/> Standardized clinical instrument, Name:																		
<input type="checkbox"/> Electronic Clinical Registry, Name:	<input type="checkbox"/> Standardized patient survey, Name:																		
<input checked="" type="checkbox"/> Electronic Claims	<input type="checkbox"/> Standardized clinician survey, Name:																		
<input type="checkbox"/> Electronic Pharmacy data	<input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period.																		
<input type="checkbox"/> Electronic Lab data																			
<input type="checkbox"/> Electronic source - other, Describe:																			
12 (2a)	<p>Sampling <i>If measure is based on a sample, provide instructions and guidance on sample size.</i></p> <p>Minimum sample size: h1</p> <p>Instructions:</p>																		
13 (2a)	<p>Type of Measure: Outcome ▶ If "Other", please describe:</p> <p>▶ If part of a composite or paired with another measure, please identify composite or paired measure</p> <p>While the measure to two types of information components, the results are not a composite as is defined by NQF, but rather a reliability adjusted measure of survival.</p> <p>Volume is used to create a volume predicted mortality for the hospital--this component of the measure is used to create greater reliability for low-volume hospitals. In the modeling for this measure, the volume predicted mortality and the observed mortality are weighted. In the model, lower volume hospitals have a higher weight on the volume predicted mortality versus the observed mortality. The opposite is true for high volume hospitals, which have a higher weight on the observed mortality. This methodology results in a reliability adjusted survival predictor.</p>																		
14 (2a)	<p>Unit of Measurement/Analysis <i>(Who or what is being measured) Check all that apply.</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Can be measured at all levels</td> <td><input type="checkbox"/> Integrated delivery system</td> </tr> <tr> <td><input type="checkbox"/> Individual clinician (e.g., physician, nurse)</td> <td><input checked="" type="checkbox"/> Health plan</td> </tr> <tr> <td><input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)</td> <td><input checked="" type="checkbox"/> Community/Population</td> </tr> <tr> <td><input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)</td> <td><input type="checkbox"/> Other <i>(Please describe):</i></td> </tr> </table>	<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system	<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input checked="" type="checkbox"/> Health plan	<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input checked="" type="checkbox"/> Community/Population	<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other <i>(Please describe):</i>										
<input type="checkbox"/> Can be measured at all levels	<input type="checkbox"/> Integrated delivery system																		
<input type="checkbox"/> Individual clinician (e.g., physician, nurse)	<input checked="" type="checkbox"/> Health plan																		
<input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice)	<input checked="" type="checkbox"/> Community/Population																		
<input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home)	<input type="checkbox"/> Other <i>(Please describe):</i>																		
15 (2a)	<p>Applicable Care Settings <i>Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Can be used in all healthcare settings</td> <td><input type="checkbox"/> Hospice</td> </tr> <tr> <td><input type="checkbox"/> Ambulatory Care (office/clinic)</td> <td><input checked="" type="checkbox"/> Hospital</td> </tr> <tr> <td><input type="checkbox"/> Behavioral Healthcare</td> <td><input type="checkbox"/> Long term acute care hospital</td> </tr> <tr> <td><input type="checkbox"/> Community Healthcare</td> <td><input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)</td> </tr> <tr> <td><input type="checkbox"/> Dialysis Facility</td> <td><input type="checkbox"/> Prescription Drug Plan</td> </tr> <tr> <td><input type="checkbox"/> Emergency Department</td> <td><input type="checkbox"/> Rehabilitation Facility</td> </tr> <tr> <td><input type="checkbox"/> EMS emergency medical services</td> <td><input type="checkbox"/> Substance Use Treatment Program/Center</td> </tr> <tr> <td><input type="checkbox"/> Health Plan</td> <td><input type="checkbox"/> Other <i>(Please describe):</i></td> </tr> <tr> <td><input type="checkbox"/> Home Health</td> <td></td> </tr> </table>	<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice	<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital	<input type="checkbox"/> Behavioral Healthcare	<input type="checkbox"/> Long term acute care hospital	<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)	<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan	<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility	<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center	<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other <i>(Please describe):</i>	<input type="checkbox"/> Home Health	
<input type="checkbox"/> Can be used in all healthcare settings	<input type="checkbox"/> Hospice																		
<input type="checkbox"/> Ambulatory Care (office/clinic)	<input checked="" type="checkbox"/> Hospital																		
<input type="checkbox"/> Behavioral Healthcare	<input type="checkbox"/> Long term acute care hospital																		
<input type="checkbox"/> Community Healthcare	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF)																		
<input type="checkbox"/> Dialysis Facility	<input type="checkbox"/> Prescription Drug Plan																		
<input type="checkbox"/> Emergency Department	<input type="checkbox"/> Rehabilitation Facility																		
<input type="checkbox"/> EMS emergency medical services	<input type="checkbox"/> Substance Use Treatment Program/Center																		
<input type="checkbox"/> Health Plan	<input type="checkbox"/> Other <i>(Please describe):</i>																		
<input type="checkbox"/> Home Health																			
IMPORTANCE TO MEASURE AND REPORT																			
<p>Note: This is a threshold criterion. If a measure is not judged to be sufficiently important to measure and report, it will not be evaluated against the remaining criteria.</p>																			

16 (1a)	<p>Is measure related to a National Priority Partners priority area? Safety reliability (for NQF staff use) Does measure address a specific NPP goal? (www.qualityforum.org/about/NPP/):</p>
17 (1a)	<p>Does the measure address a high impact aspect of healthcare patient/societal consequences of poor quality</p> <p>Summary of Evidence: This measure addresses mortality in an extremely high risk procedure (esophagectomy) and is an outcome measure which is of interest to both consumers and purchasers.</p> <p>While this is a low volume procedure, it is one that has great variation in mortality across hospitals. The absolute difference between low and high volume hospital mortality exceeded 10% (Birkmeyer, et al. As indicated, mortality in US hospitals varies for esophagectomy surgeries--there are significant documented differences between high and low performing hospitals [4]. Higher volumes are associated with better outcomes including lower mortality.</p> <p>This measure improves upon the technology of surgical mortality measurement. It overcomes three problems with existing mortality measures: 1) Mortality rates are often too "noisy" to reflect hospital quality with surgery (particularly among lower volume hospitals), 2) volume alone is a weak proxy for most procedures, and 3) when both volume and mortality are reported as separate indicators it is difficult to understand which measure is more important. [1] Given the relatively small number of esophagectomy procedures performed annually in the United States, it is important that a mortality measure is designed to reliably measure low volume hospitals, and this measure specifically addresses hospitals which perform relatively few procedures. Up to this point, in order to measure this outcome, other measure developers have added less significant procedures to the denominator in order to gain reliability. Yet, consumers and purchasers would benefit more from knowing specifically where to get one of the most risky procedures performed. The information from the survival predictor is more reliable for these small volume counts than existing measures. In addition, this measure can be applied to the nation, states, or regions.</p> <p>Birkmeyer and Dimick (2009)[4] show that differences in mortality can be predicted using a reliability adjusted mortality rate (a weighted combination of volume and mortality) which is particularly relevant for selective-referral or public reporting contexts. They reduce the effects of random chance (statistical noise) and as a result with CABG, for example, more than half of the observed variation can be attributed to statistical noise. When they sorted hospitals simply on actual (risk-adjusted) mortality, rates varied from 1.4% to 11.0% across hospital quintiles (Figure 1 in White Paper [1]). After they adjusted for reliability, however, the mortality rates varied considerably less, from 3.3% to 6.3%. Although the almost twofold variation in mortality still suggests ample opportunity for quality improvement, these data underscore the importance of accounting for chance in understanding variation in hospital outcomes.</p> <p>Citations² for Evidence:</p> <p>[1] DeFrances, C.J., Lucas, CA, Bule, VC., Golosinskiy, A. 2006 National Hospital Discharge Survey, National health statistics reports, no. 5. Hyattsville, MD: National Center for Health Statistics. 2008. Accessed on 12/17/08 at http://www.cdc.gov/nchs/data/nhsr/nhsr005.pdf</p> <p>[2] The National Hospital Bill: The Most Expensive Conditions by Payer, 2006. Statistical Brief #59. File accessed on March 16, 2009, at: http://www.hcup-us.ahrq.gov/reports/statbriefs/sb59.jsp Produced by AHRQ, Center for Delivery, Organization, and Markets, Healthcare Cost and Utilization Project, Nationwide Inpatient Sample, 2006.</p> <p>[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf</p>

² Citations can include, but are not limited to journal articles, reports, web pages (URLs).
NQF Measure Submission Form, V3.1

	<p>[4] Birkmeyer, J.D., and Dimick, J.B. (2009) Understanding and reducing variation in surgical mortality. <i>Annu. Rev. Med.</i> 2009. 60:405-15.</p>
<p>18 (1b)</p>	<p>Opportunity for Improvement <i>Provide evidence that demonstrates considerable variation, or overall poor performance, across providers.</i></p> <p>Summary of Evidence: In 2002, a systematic review of the literature on the volume-outcome relationship found that there was a significant relationship between hospital volume and outcomes for esophagectomy surgery. Unlike the relationship for CABG which was less robust, both surgical procedures for esophagectomy and pancreatectomy were robust. [5]</p> <p>In a 2003 article on surgical volume and quality of care, Dimick, Pronovost, et al, found that high volume hospitals had a mortality rate of 2.5% in Maryland hospitals (using data from 1994-1998) while low volume hospitals mortality rate was 15.4% (p<0.001) with a case-mixed adjusted odds ratio of death equal to 5.7 (95% CI, 2.0-16; p=<0.001). [5a]</p> <p>Given the findings related to volume of procedures, Silber et al., [7] explored the relative contribution of complication rates and failure to rescue rates to mortality and found that complication rates were more likely influenced by patient factors while failure to rescue rates of those with complications was more related to hospital factors. Thus, it may be that higher volume hospitals are better at rescuing patients with complications. Silbers finding, in conjunction with the volume information, suggests lower volume hospitals with worse mortality rates could in fact address this through better care following the procedure, thereby reducing their overall rate.</p> <p>Note: Birkmeyer and Dimmick [4] indicate it is also likely that some lower volume hospitals would also have lower mortality rates.</p> <p>Citations for Evidence: [5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. <i>Annals of Internal Medicine</i>, Sept 1;137(6):511-20</p> <p>[5a] Dimick, J.B., Pronovost, P.J., Cowan, J.A., and Lipsett, P.A. (2003). Surgical volume and quality of care for esophageal resection: do high volume hospitals have fewer complications? <i>Ann Thorac Surg.</i>, 75:337-341.</p> <p>[6] Epstein, A.J., Rathore, S.S., Krumholz, H.M., and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>[7] Silber, J.H., Rosenbaum, P.R., Trudeau, M.E., et al. 2005. Changes in prognosis after the first postoperative complication. <i>Medical Care</i>, 43:122-31.</p>
<p>19 (1b)</p>	<p>Disparities <i>Provide evidence that demonstrates disparity in care/outcomes related to the measure focus among populations.</i></p> <p>Summary of Evidence: It is more likely that minorities will be treated at a low volume facility, and as a result are likely to be impacted by higher mortality rates. In an analysis of the National Inpatient Sample, Epstein, Rathore and Krumholz (2005)[6] found that a greater proportion of patients treated in low volume hospitals for CABG and PCI were non-white, while a lower proportion of non-white patients presented as "elective" admissions or patients received in transfer as compared to patients in high volume hospitals. We expect that there would be similar findings for esophagectomy surgery.</p> <p>In the survival predictor, the denominator and numerator are restricted to elective procedures, therefore, it is anticipated there may be a smaller non-white, and low SES population in the denominator and numerator.</p>

	<p>Citations for evidence:</p> <p>[4, p. 3-5]</p>
<p>20 (1c)</p>	<p>If measuring an Outcome Describe relevance to the national health goal/priority, condition, population, and/or care being addressed: An esophageal surgical procedure is a high risk procedure, and a very expensive procedure, and only limited information is available nationally on the risk of mortality associated with the procedure. Other entities with clinical information are not publicly reporting mortality rates of esophageal procedures by hospital provider. This measure is designed to give feedback to hospitals across the country as well as to provide information for decision-making by consumers and purchasers.</p> <p>Mortality in US hospitals varies for esophageal surgeries--there are documented differences between high and low performing hospitals [4,5a]. Higher volumes are associated with better outcomes including lower mortality. In addition to being a high risk surgery, this surgery is one of the high cost procedures.</p> <p>This measure is highly relevant to both consumers and purchasers, given its high cost both in terms of lives lost and dollars spent. National purchasers are interested in comparative information on hospitals nationwide. Pauly (1996) in a study of purchaser interests in hospital performance reporting found that mortality ratings were more important to purchasers than were morbidity or complications. [9] Health plans are interested in contracting with centers of excellence, which can be identified through the results of survival predictor in combination with other information on cost and quality. Consumers have shown their interest in other surgical mortality by requesting reports from the state of Pennsylvania [10]; an earlier study by IOM (Lohr, Donaldson and Walker 1991) found that consumers were interested in hospital mortality rates, but did not perceive this information to be available.[11] Hibbard and Jewett found that consumers were more interested in "undesirable events" (such as mortality, complications, infections) than in "desirable events." [12]</p> <p>[5a] Dimick, J.B., Pronovost, P.J., Cowan, J.A., and Lipsett, P.A. (2003). Surgical volume and quality of care for esophageal resection: do high volume hospitals have fewer complications? <i>Ann Thorac Surg.</i>, 75:337-341.</p> <p>[9] Pauly, M.V., Brailer, D.J.Kroch, E., and Even-Shoshan, O. Measuring Hospital Outcomes from a Buyer's Perspective. <i>American Journal of Medical Quality</i>, 11(8): Fall 1996.</p> <p>[10] Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.</p> <p>[11]Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. <i>Quality Review Bulletin</i> 17:242-53.</p> <p>[12]Hibbard, J.H. and Jewett, J.(1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? <i>Medical Care Research and Review.</i>, Vol 53(1): 28-47.</p> <p>If not measuring an outcome, provide evidence supporting this measure topic and grade the strength of the evidence <i>Summarize the evidence (including citations to source) supporting the focus of the measure as follows:</i></p> <ul style="list-style-type: none"> • <u>Intermediate outcome</u> - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit. • <u>Process</u> - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s). • <u>Structure</u> - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.

	<ul style="list-style-type: none"> • <u>Patient experience</u> - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public. • <u>Access</u> - evidence that an association exists between access to a health service and the outcomes of, or experience with, care. • <u>Efficiency</u>- demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality. <p>Type of Evidence <i>Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Evidence-based guideline</td> <td><input checked="" type="checkbox"/> Quantitative research studies</td> </tr> <tr> <td><input checked="" type="checkbox"/> Meta-analysis</td> <td><input type="checkbox"/> Qualitative research studies</td> </tr> <tr> <td><input checked="" type="checkbox"/> Systematic synthesis of research</td> <td><input type="checkbox"/> Other (<i>Please describe</i>):</td> </tr> </table> <p>Overall Grade for Strength of the Evidence³ (<i>Use the USPSTF system, or if different, also describe how it relates to the USPSTF system</i>): Moderate</p> <p>Summary of Evidence (<i>provide guideline information below</i>): Over 100 articles published related to volume and outcome relationship, with some inconsistency in results. Systematic review of the literature conducted in 2002. No review since that time.</p> <p>Citations for Evidence:</p> <p>[5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. <i>Annals of Internal Medicine</i>, Sept 1;137(6):511-20.</p> <p>[5a] Dimick, J.B., Pronovost, P.J., Cowan, J.A., and Lipsett, P.A. (2003). Surgical volume and quality of care for esophageal resection: do high volume hospitals have fewer complications? <i>Ann Thorac Surg.</i>, 75:337-341.</p> <p>[14] Birkmeyer, J.D., Dimick, J.B., Staiger, D.O. (2006) Operative mortality and procedure volume as predictors of subsequent hospital performance. <i>Ann Surg.</i> 243:411-417.</p> <p>[15] Dimick, JB, Welch HG, Birkmeyer JD. (2004) Surgical mortality as an indicator of hospital quality: The problem with small sample size. <i>JAMA</i>, 292:847-851.</p> <p>[4] Birkmeyer, JD., and Dimick, JB. (2009) Understanding and reducing variation in surgical mortality. <i>Annu. Rev. Med.</i> 60:405-15.</p> <p>[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. <i>Medical Care</i>, 47(2): 226-233.</p> <p>[18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. <i>N Engl J Med.</i> 1979;301:1364-9.</p>	<input type="checkbox"/> Evidence-based guideline	<input checked="" type="checkbox"/> Quantitative research studies	<input checked="" type="checkbox"/> Meta-analysis	<input type="checkbox"/> Qualitative research studies	<input checked="" type="checkbox"/> Systematic synthesis of research	<input type="checkbox"/> Other (<i>Please describe</i>):
<input type="checkbox"/> Evidence-based guideline	<input checked="" type="checkbox"/> Quantitative research studies						
<input checked="" type="checkbox"/> Meta-analysis	<input type="checkbox"/> Qualitative research studies						
<input checked="" type="checkbox"/> Systematic synthesis of research	<input type="checkbox"/> Other (<i>Please describe</i>):						
<p>21 (1c)</p>	<p>Clinical Practice Guideline <i>Cite the guideline reference; quote the specific guideline recommendation related to the measure and the guideline author's assessment of the strength of the evidence; and summarize the rationale for using this guideline over others.</i></p> <p>Guideline Citation:</p>						

³The strength of the body of evidence for the specific measure focus should be systematically assessed and rated, e.g., USPSTF grading system www.ahrq.gov/clinic/uspstmeth.htm: **A** - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. **B** - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. **C** - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. **D** - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. **I** - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

	<p>Specific guideline recommendation:</p> <p>Guideline author’s rating of strength of evidence (<i>If different from USPSTF, also describe it and how it relates to USPSTF</i>):</p> <p>Rationale for using this guideline over others:</p>
<p>22 (1c)</p>	<p>Controversy/Contradictory Evidence Summarize any areas of controversy, contradictory evidence, or contradictory guidelines and provide citations.</p> <p>Summary: There are three areas of possible contention with this measure--</p> <p>1) The volume-outcome relationship has been questioned for some procedures [6, 17]</p> <p>More than 100 studies have demonstrated better results at high-volume hospitals with cardiovascular surgery, major cancer resections (esophagectomy), and other high-risk procedures.[18, 20]</p> <p>All studies listed here were done to determine whether there was a volume-outcome relationship for hospitals performing surgical procedures--they all documented that there were differences between low volume hospitals and high volume hospitals in mortality--and the evidence for this relationship appears strongest for two procedures, esophagectomy and pancreatectomy with high volume hospitals having less mortality. In the case of esophagectomy the risk of dying was 4-fold more at low volume hospitals. [5a]</p> <p>2) That outcome measures must be risk-adjusted unless there is evidence to show it is not needed (NQF). The survival predictor measure predicts better than volume or mortality alone, and is as good a predictor as risk-adjusted mortality. When testing the unadjusted survival predictor against risk-adjusted mortality there was a (.96) correlation. [4] See Section 28 of this form for details.</p> <p>3) The weighting of input measures into composites. Existing approaches rely on overly simplistic approaches. Among these, assigning equal weight to all measures (i.e., the all or none approach) and relying on expert opinion are the most common. The survival predictor relies on empiric methods for weighting the input measures.</p> <p>Citations: [6] Andrew J Epstein, Saif S Rathore, Harlan M Krumholz and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. BMC Health Services Research, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>[5a] Dimick, J.B., Pronovost, P.J., Cowan, J.A., and Lipsett, P.A. (2003). Surgical volume and quality of care for esophageal resection: do high volume hospitals have fewer complications? Ann Thorac Surg., 75:337-341.</p> <p>[16]Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. Medical Care, 47(2): p. 232.</p> <p>[17] Edward L. Hannan, PhD; Chuntao Wu, PhD; Thomas J. Ryan, MD; Edward Bennett, MD; Alfred T. Culliford, MD; Jeffrey P. Gold, MD; Alan Hartman, MD; O. Wayne Isom, MD; Robert H. Jones, MD; Barbara McNeil, MD, PhD; Eric A. Rose, MD; Valavanur A. Subramanian, MD. Do Hospitals and Surgeons With Higher Coronary Artery Bypass Graft Surgery Volumes Still Have Lower Risk-Adjusted Mortality Rates? Circulation. 2003;108:795-801.</p> <p>[18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. N Engl J Med. 1979;301:1364-9.</p> <p>[20] Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. JAMA. 1998;280:1747-51.)</p>
<p>23</p>	<p>Briefly describe how this measure (as specified) will facilitate significant gains in healthcare quality</p>

(1)	<p>related to the specific priority goals and quality problems identified above: This measure of predicted survival improves upon the reliability of mortality results for high risk surgical procedures, such as esophagectomy. For the first time, this measure produces reliable mortality/survivability information on smaller volume hospitals, as well as high volume hospitals. Hospitals across the country will have information available through voluntary public reporting.</p>
SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
<p>Note: Testing and results should be summarized in this form. However, additional detail and reports may be submitted as supplemental information or provided as a web page URL. If a measure has not been tested, it is only potentially eligible for time-limited endorsement.</p>	
24	<p>Supplemental Testing Information: attached <input checked="" type="checkbox"/> OR Web page URL:</p>
25	<p>Reliability Testing</p> <p>(2b) Data/sample: Data was a 100% sample from the Medicare Analysis Provider and Review (MEDPAR) files for 2000-2003, these files contain 100% of Medicare hospitalizations for years specified. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the esophageal resection surgery. We selected only those resections where there was an esophageal cancer diagnosis present, thereby creating a more homogenous risk pool [3].</p> <p>Note: Needleman, Buerhaus, et al., (2003) concluded after applying operational tests on Medicare data for adverse outcomes and all-patient hospital data from 11 states, that Medicare data could be used to assess quality in hospitals.[20] Given the lack of a national all-patient database, MEDPAR data was used in development and testing of the models.</p> <p>Analytic Method:</p> <p>Model Development</p> <p>We used an empirical Bayes approach to combine mortality rates with information on hospital volume at each hospital. In traditional empirical Bayes methods, a point estimate (e.g., mortality rate observed at a hospital) is adjusted for reliability by shrinking it towards the overall mean (e.g., overall mortality rate in the population) [21,22]. We modified this traditional approach by shrinking the observed mortality rate back toward the mortality rate expected given the volume at that hospital—we refer to this as the “volume-predicted mortality” (See attached White Paper TECHNICAL APPENDIX for the mathematical details of this method). With this approach, the observed mortality rate is weighted according to how reliably it is estimated, with the remaining weight placed on the information regarding hospital volume. Because this method includes observed data to the extent that it is useful, and only relies on the proxy measure to the extent necessary, it ensures an optimal combination of these two quality domains. [3]</p> <p>The two inputs to the survival predictor measure are mortality rates and procedure volume for each of the six included operations. Procedure-specific mortality rates were calculated for all hospitals over a 2-year period (2000-01) and this was used as the first input. Hospital volume was calculated as the number of Medicare cases performed during the same time period. For each operation, the relationship between hospital volume and risk-adjusted mortality was modeled using linear regression. (Details of the risk-adjustment strategy will be discussed below.) After testing the fit of several transformations, hospital volume was modeled as the natural log of the continuous volume variable, which is the same approach used in our previous work [23]. Using this regression model, we estimated the volume-predicted mortality, the second input to the survival predictor measure.</p>

	<p>We then used the empirical Bayes approach to create an optimal combination of these two inputs. This survival predictor measure theoretically provides the best estimate of a hospital's true mortality rate, taking into account the both available inputs [21,22]. The combined survival predictor measure was calculated as follows: mortality prediction = (weight)*(observed mortality) + (1-weight)*(volume-predicted mortality). The weight placed on the point estimate of mortality is the reliability, or ratio of signal to signal plus noise, calculated as follows: weight = variation among hospitals/(variation among hospitals + variation within hospitals). The variation among hospitals was calculated as the variance in observed mortality rates for the hospitals included in the sample. The variation within hospitals was calculated as the standard error of the mortality rate at each hospital. With this method, more weight is placed on the observed mortality rate when a hospital has a high number of cases because it is estimated with more reliability; less weight is placed on the observed mortality rate when a hospital performs a low number of cases because of its lower reliability.</p> <p>A calculation worksheet with examples is attached.</p> <p>Testing Results: Hospital caseloads and the weights applied to each input to the survival predictor measure varied for each procedure studied (see Table 1 in white paper [3]). For esophageal surgical procedures, a procedure with relatively low hospital caseloads, the weight applied to the volume input was .86 ([3]-Table 1), indicating that the observed mortality was less reliable than the volume predicted mortality.</p> <p>The survival predictor (mortality) measure explained a large proportion of non-random, hospital-level variation in risk-adjusted mortality rates (see Table 2 in White Paper [3]). For esophageal procedures, the survival predictor explained 44% of the hospital level variation in mortality rates; this compares to 14% for observed mortality and 33% for volume of esophageal surgeries. Measures with low reliability or correlation explain little variation. The correlation between the survival predictor and risk-adjusted mortality was (.96) ([16] p. 232), and the amount of variation explained was 44% [3]. This is an adequate level of reliability.</p> <p>Citations:</p> <p>[3] p. 19 (Table 2)</p> <p>[21] Morris CN. Parametric Empirical Bayes Inference: Theory and Applications. J Am Stat Assoc 1988;78:47-55.</p> <p>[22] McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) Frontiers in Health Policy Research. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.</p> <p>[23] Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. N Engl J Med. 2003;349:2117-2127.</p> <p>[20] Needleman, J., Buerhaus, P.I., Mattke, S., Stewart, M., and Zelevinsky, M. (2003). Health Services Research 38.6, Part I; 1487-1508.</p> <p>[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. Medical Care, 47(2): 226-233.</p>
<p>26</p> <p>(2c)</p>	<p>Validity Testing</p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p>

Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the esophagectomy surgery. We only included patients with cancer as a diagnosis (which is the bulk of the population and highly correlated with the full population).

Analytic Method: We determined the value of our survival predictor (mortality) measure by establishing whether it explained hospital-level variation in risk-adjusted mortality rates and by assessing to what degree it was able to predict future hospital performance. We first estimated the proportion of variation in hospital-level mortality (2000-01) explained by the survival predictor measure using random effects logistic regression models. For these analyses, we estimated the proportional change in the hospital-level variance in mortality rates, which was determined from the standard deviation of the random effect, after adding each measure to the model [14,22].

We next compared the ability of the survival predictor measure to the individual measures, mortality rates and hospital volume. We should note that these analyses focus on explaining systematic, or non-random, variation, since measurement error (random error) is accounted for and subtracted from the total variation in all analyses [22,24].

We next determined the extent to which the composite measure predicts future risk-adjusted mortality. For this analysis, hospitals were ranked based on each measure from the earlier time period (data from years 2000-01) and divided into four equal size groups (quartiles at the patient level). The subsequent risk-adjusted mortality rates for each quartile of performance were then calculated (data from years 2002-03). We present the subsequent mortality rates across quartiles of the esophagectomy survival predictor measure to graphically demonstrate its usefulness in discriminating among hospitals for the entire spectrum of performance. To compare the predictive ability of the composite measures and individual measures, we also present the subsequent mortality rates in the "worst" compared to the "best" quartile in the White Paper ([3], p. 22) "Quartiles of Performance Measures (2000-2001). This table reflects how well the unadjusted survival predictor created on 2000-2001 data compares to risk-adjusted mortality in 2002-2003 data.

Note: The risk-adjusted mortality rate for esophagectomy was constructed using standard methods. We determined the ratio of actual deaths or complications to the number of expected deaths (the O/E ratio). The number of expected deaths was the sum over all patients of the predicted probability of death or complications derived from a logistic regression model estimated on all patients undergoing esophagectomy surgery. The dependent variable in the logistic model was death or complications and the independent variables were patient covariates. The patient characteristics included age, gender, race, admission acuity, and co-existing diseases using the Elixhauser method. A zip code level measure of socio-economic status was derived from 2000 census data.

Testing Results: While some measures are good at discriminating top performers or bottom performers, this measure is good at prediction across entire spectrum of performance. [See White paper [3]: Figures p. 21-22) for a graphical demonstration of the usefulness of the survival predictor in discriminating among hospitals across the entire spectrum of performance.] To compare the predictive ability of the reliability adjusted survival predictor versus the individual components (volume and observed mortality) we also present the subsequent mortality rates in the "worst" compared to the "best" quartile.

[22]. McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) Frontiers in Health Policy Research. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.

[14] Birkmeyer JD, Dimick JB, Staiger DO. Operative mortality and procedure volume as predictors of subsequent hospital performance. *Ann Surg* 2006;243:411-417.

[24] Zaslavsky AM, Cleary PD. Dimensions of plan performance for sick and healthy members on the Consumer Assessments of Health Plans Study 2.0 survey. *Med Care* 2002;40:951-964.

<p>27 (2d)</p>	<p>Measure Exclusions <i>Provide evidence to justify exclusion(s) and analysis of impact on measure results during testing.</i></p> <p>Summary of Evidence supporting exclusion(s): The developers defined the denominator to minimize potential for case mix differences between hospitals, they created homogenous sub-groups, in this case only those esophageal resections with a cancer diagnosis.</p> <p>Only those hospitals with elective cases will have a survival predictor, since the primary goal of the measure is to provide information for selection of a specific hospital for the esophagectomy procedure.</p> <p>Citations for Evidence:</p> <p>Data/sample:</p> <p>Analytic Method:</p> <p>Testing Results:</p>
<p>28 (2e)</p>	<p>Risk Adjustment Testing <i>Summarize the testing used to determine the need (or no need) for risk adjustment and the statistical performance of the risk adjustment method.</i></p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the esophagectomy surgery. We created homogenous patient subgroups, including those with diagnosis codes indicating that the patient had an esophageal cancer.</p> <p>Analytic Method: Sensitivity analysis. We performed a sensitivity analysis to determine whether risk-adjustment of the mortality input was important in improving the predictive ability of the survival predictor measure. Risk-adjustment was performed using logistic regression to estimate expected mortality rates for each hospital based on patient age, gender, race, urgency of operation, median income, and coexisting diseases. Coexisting diseases were determined from secondary diagnostic codes using the methods of Elixhauser (16). The observed mortality rate at each hospital was then divided by the expected mortality rate to yield the ratio of observed/expected deaths (O/E ratio). The O/E ratio was multiplied by the average mortality rate for each operation to yield a risk-adjusted mortality rate. To determine the value of risk-adjustment in the context of selective referral, we compared the ability of risk-adjusted and unadjusted composite measures to predict subsequent performance.</p> <p>Testing Results: In sensitivity analysis, composite measures based on an unadjusted mortality input and a risk-adjusted mortality input had a correlation of (.95) and thus were equally good at predicting future performance (See pages 21-22 in the White Paper [3]).</p> <p>► If outcome or resource use measure not risk adjusted, provide rationale: Because risk-adjusted mortality is not available publicly except for limited locations, the capacity to use unadjusted mortality is very desirable, especially since it was shown to provide (under this methodology) an equal result. This measure will allow measurement to occur across the United States, providing information to national companies, health plans and consumers.</p>
<p>29 (2g)</p>	<p>Testing comparability of results when more than 1 data method is specified (e.g., administrative claims or chart abstraction)</p> <p>Data/sample: not applicable</p> <p>Analytic Method:</p>

	Results:
30 (2f)	<p>Provide Measure Results from Testing or Current Use Results from testing</p> <p>Data/sample: same as described above, results for survival predictor in White Paper [3]available on Website and Validation results for composite in [16] Staiger, Dimick et al., Medical Care 2009</p> <p>Methods to identify statistically significant and practically/meaningfully differences in performance: Bayesian Hierarchical methods using new shrinkage estimator Empirical Bayesian methods to determine weights Correlations Calculated the amount of variation predicted by survival predictor as a percentage of all hospital-level variation (adjusted for sampling variation)--analogous to a R-squared from a regression that summarizes the ability of the predictor to explain the hospital level variation in mortality for esophageal resection surgery. Predictor was tested against the "gold standard" --risk adjusted mortality</p> <p>Results: See White Paper [3]</p>
31 (2h)	<p>Identification of Disparities</p> <p>► If measure is stratified by factors related to disparities (i.e. race/ethnicity, primary language, gender, SES, health literacy), provide stratified results:</p> <p>► If disparities have been reported/identified, but measure is not specified to detect disparities, provide rationale: .</p>
USABILITY	
32 (3)	<p>Current Use Testing completed If in use, how widely used Nationally ► If "other," please describe: Survival Predictor for Pancreatectomy and Esophagectomy in use--see URL.</p> <p><input checked="" type="checkbox"/> Used in a public reporting initiative, name of initiative: Leapfrog Hospital Survey Sample report attached <input type="checkbox"/> OR Web page URL: https://www.leapfroggroup.org/cp</p>
33 (3a)	<p>Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)</p> <p>Data/sample:</p> <p>Methods:</p> <p>Results: See following citations reflecting consumer use of mortality information:</p> <p>[10]Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.</p> <p>[11]Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. Quality Review Bulletin 17:242-53.</p> <p>[12]Hibbard, J.H. and Jewett, J.(1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? Medical Care Research and Review., Vol 53(1): 28-47.</p>
34 (3b, 3c)	<p>Relation to other NQF-endorsed™ measures</p> <p>► Is this measure similar or related to measure(s) already endorsed by NQF (on the same topic or the same target population)? <i>Measures can be found at www.qualityforum.org under Core Documents.</i></p> <p><i>Check all that apply</i></p> <p><input type="checkbox"/> Have not looked at other NQF measures <input checked="" type="checkbox"/> Other measure(s) on same topic <input checked="" type="checkbox"/> Other measure(s) for same target population <input type="checkbox"/> No similar or related measures</p>

	<p>Name and number of similar or related NQF-endorsed™ measure(s): Esophageal Resection Volume IQI 1, NQF endorsed measure 0361; Esophageal Resection Mortality Rate (IQI 8) NQF # 0360.</p> <p>Are the measure specifications harmonized with existing NQF-endorsed™ measures? Not harmonized ▶ If not fully harmonized, provide rationale: The AHRQ measures include many less difficult procedures than does the Survival Predictor. It is unlikely that the mortality rates for the AHRQ measures are as high given the smoothing effect of adding procedures to the specifications that do not require as much surgical skill as the principal codes in the Survival Predictor. Additionally, the Survival Predictor is based on elective procedures, with a cancer diagnosis. Patients with cancer could elect where to have the procedure performed.</p> <p>Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: This measure provides the ability to produce reliable mortality results for low volume hospitals, other measures do not have this capacity. In addition, the access to data nationally for other esophagectomy mortality measures does not exist.</p>
FEASIBILITY	
35 (4a)	<p>How are the required data elements generated? <i>Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data elements are generated concurrent with and as a byproduct of care processes during care delivery (e.g., blood pressure or other assessment recorded by personnel conducting the assessment)</p> <p><input type="checkbox"/> Data elements are generated from a patient survey (e.g., CAHPS)</p> <p><input checked="" type="checkbox"/> Data elements are generated through coding performed by someone other than the person who obtained the original information (e.g., DRG or ICD-9 coding on claims)</p> <p><input checked="" type="checkbox"/> Other, Please describe: Data are currently submitted to Leapfrog via a secure online survey-</p>
36 (4b)	<p>Electronic Sources All data elements</p> <p>▶ If all data elements are not in electronic sources, specify the near-term path to electronic collection by most providers:</p> <p>▶ Specify the data elements for the electronic health record: volume of esophageal resection procedures, observed death during inpatient stay, related to esophageal resection cases with a cancer diagnosis</p>
37 (4c)	<p>Do the specified exclusions require additional data sources beyond what is required for the other specifications? No</p> <p>▶ If yes, provide justification:</p>
38 (4d)	<p>Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure: It is unlikely that this procedure, or inpatient death will be inaccurately coded or not coded given the high cost of procedure and the accompanying death.</p> <p>Describe how could these potential problems be audited: If problems were identified, a chart review of cases could be performed.</p> <p>Did you audit for these potential problems during testing? No If yes, provide results:</p>
39 (4e)	<p>Testing feasibility Describe what have you learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: Initial results only available for Esophagectomy, Pancreatectomy. CABG will be released in 2009</p>
CONTACT INFORMATION	
40	<p>Web Page URL for Measure Information Describe where users (implementers) should go for more details on specifications of measures, or assistance in implementing the measure. Web page URL: https://leapfrog.medstat.com for access to Survival Predictor White Paper</p>
41	<p>Measure Steward Point of Contact</p>

	<p>First Name: MI: Last Name: Credentials (MD, MPH, etc.): Organization: The Leapfrog Group % The Academy Street Address: 1150 17th St., NW, Suite 600 City: Washington State: DC ZIP: 20036 Email: Telephone: ext:</p>
42	<p>Measure Developer Point of Contact If different from Measure Steward First Name: Justin MI: B Last Name: Dimick Credentials (MD, MPH, etc.): MD, MPH Organization: Department of Surgery, University of Michigan, M-SCORE offices, Suite 201 and 202 Street Address: 211 N. Fourth Avenue City: Ann Arbor State: MI ZIP: 48104 Email: jdimick@umich.edu Telephone: ext:</p>
ADDITIONAL INFORMATION	
43	<p>Workgroup/Expert Panel involved in measure development Workgroup/panel used ► If workgroup used, describe the members' role in measure development: Research team led by Justin Dimick, MD, MPH; ► Provide a list of workgroup/panel members' names and organizations:</p> <p>Douglas Staiger Ph.D., Department of Economics and the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth College, Hanover, New Hampshire</p> <p>John D. Birkmeyer, MD Michigan Surgical Collaborative for Outcomes Research and Evaluation Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Onur Baser, Ph.D. Michigan Surgical Collaborative for Outcomes Research and Evaluation Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Research supported by the National Institute on Aging</p>
44	<p>Measure Developer/Steward Updates and Ongoing Maintenance Year the measure was first released: 2008 Month and Year of most recent revision: August 2008 What is the frequency for review/update of this measure? Annual When is the next scheduled review/update for this measure? New coefficients for August 2009</p>
45	<p>Copyright statement/disclaimers: none</p>
46	<p>Additional Information: All measure information is available at https://leapfrog.medstat.com Please contact measure developer prior to use to assure all necessary items have been accessed.</p>
47	<p>I have checked that the submission is complete and any blank fields indicate that no information is provided. <input checked="" type="checkbox"/></p>
48	<p>Date of Submission (MM/DD/YY): Revised submission dated 3/18/09</p>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

The measure information you submit will be shared with NQF’s Steering Committees and Technical Advisory Panels to evaluate measures against the NQF criteria of importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Four conditions (as indicated below) must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, *importance to measure and report*, in order to be evaluated against the remaining criteria. References to the specific measure evaluation criteria are provided in parentheses following the item numbers. Please refer to the *Measure Evaluation Criteria* for more information at www.qualityforum.org under Core Documents. Additional guidance is being developed and when available will be posted on the NQF website.

Use the tab or arrow (↓→) keys to move the cursor to the next field (or back ←↑). There are three types of response fields:

- drop-down menus - select one response;
- check boxes - check as many as apply; and
- text fields - you can copy and paste text into these fields or enter text; these fields are not limited in size, but in most cases, we ask that you summarize the requested information.

Please note that URL hyperlinks do not work in the form; you will need to type them into your web browser.

Be sure to answer all questions. Fields that are left blank will be interpreted as no or none. **Information must be provided in this form.** Attachments are not allowed except to provide additional detail or source documents for information that is summarized in this form. If you have important information that is not addressed by the questions, they can be entered into item #46 near the end of the form.

For questions about this form, please contact the NQF Project Director listed in the corresponding call for measures.

CONDITIONS FOR CONSIDERATION BY NQF	
	Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards.
A (A)	Public domain or Measure Steward Agreement signed: Public domain - Agreement not required (If no, do not submit) <i>Template for the Measure Steward Agreement is available at www.qualityforum.org under Core Documents.</i>
B (B)	Measure steward/maintenance: Is there an identified responsible entity and process to maintain and update the measure on a schedule commensurate with clinical innovation, but at least every 3 years? Yes, information provided in contact section (If no, do not submit)
C (C)	Intended use: Does the intended use of the measure include BOTH public reporting AND quality improvement? Yes (If no, do not submit)
D (D)	Fully developed and tested: Is the measure fully developed AND tested? Yes, fully developed and tested (If not tested and no plans for testing within 24 months, do not submit)

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.1

March 2009

	(for NQF staff use) NQF Review #: HOE-024-08 NQF Project: Hospital Outcomes and Efficiency
MEASURE SPECIFICATIONS & DESCRIPTIVE INFORMATION	
1	Information current as of (date- MM/DD/YY):
2	Title of Measure: Survival Predictor for Pancreatic Resection Surgery
3	Brief description of measure ¹ : A reliability adjusted measure of pancreatic resection surgical performance that optimally combines two important domains: Pancreatic resection hospital volume and pancreatic operative mortality, to provide predictions on pancreatic survival rates for hospitals. This measure is calculated based on data from administrative claims information.
4 (2a)	<p>Numerator Statement: Note: Because of the type of modeling done for this Survival Predictor--the information is not readily split into Numerator/ Denominator statements. Thus, we describe the two domains and their coding and data needs in this section.</p> <p>The formula for calculating the survival predictor has two components, one is a volume predicted mortality rate, and the second is an observed mortality rate. The volume predicted mortality rate reflects the hospitals experience performing pancreatic resection surgeries (thus, it includes all pancreatic resection surgeries) and uses mortality for all hospitals at that specific volume to create the volume predicted mortality. The input data from the hospitals for this domain is a volume count of all pancreatic resections performed in the hospital.</p> <p>The second domain is the observed mortality, for this domain the population is narrowed to a homogenous group of pancreatic resections with a diagnosis of cancer, the data needed for this domain is the number of observed deaths occurring for pancreatic resection cases with cancer, within the inpatient setting.</p> <p>Note: All data is available in administrative claims information. In the case of Leapfrog's implementation hospitals are asked to submit aggregated information from their claims data. No personal health information is submitted to Leapfrog. Other users of the measure may have direct access to administrative data.</p> <p>Time Window: 12 months</p> <p>Numerator Details (Definitions, codes with description): For the volume predicted mortality, hospitals count the number of pancreatic resection cases using the following codes:</p> <p>ICD-9-CM Procedure Codes:</p> <p>Any pancreaticoduodenectomy:</p> <p>5251 Proximal Pancreatectomy 5253 Radical Subtot Pancreatectomy 526 Total Pancreatectomy 527 Radical Pancreatectomy</p> <p>See calculation worksheet for with examples of how volume-predicted mortality is used in the model.</p> <p>For the observed mortality domain, the hospital submits the observed deaths for pancreatic resection</p>

¹ Example of measure description: Percentage of adult patients with diabetes aged 18-75 years receiving one or more A1c test(s) per year.
NQF Measure Submission Form, V3.1

	<p>cases using the following codes:</p> <p>ICD-9-CM Procedure Codes:</p> <p>5251 Proximal Pancreatectomy 5253 Radical Subtot Pancreatectomy 526 Total Pancreatectomy 527 Radical Pancreatectomy</p> <p>And, one of the following pancreatic cancer diagnoses codes:</p> <p>1521 MALIGNANT NEOPL JEJUNUM 1522 MALIGNANT NEOPLASM ILEUM 1523 MAL NEO MECKEL'S DIVERT 1528 MAL NEO SMALL BOWEL NEC 1529 MAL NEO SMALL BOWEL NOS 1560 MALIG NEO GALLBLADDER 1561 MAL NEO EXTRAHEPAT DUCTS 1562 MAL NEO AMPULLA OF VATER 1568 MALIG NEO BILIARY NEC 1569 MALIG NEO BILIARY NOS 1570 MAL NEO PANCREAS HEAD 1571 MAL NEO PANCREAS BODY 1572 MAL NEO PANCREAS TAIL 1573 MAL NEO PANCREATIC DUCT 1574 MAL NEO ISLET LANGERHANS 1578 MALIG NEO PANCREAS NEC 1579 MALIG NEO PANCREAS NOS</p> <p>Thus, the observed mortality is based on the volume count of pancreatic resections and an actual count of deaths occurring for that subset of pancreatic resections with cancer as a diagnosis.</p> <p>See Calculation Worksheet for how the two domains are used to create the Survival Predictor.</p>
<p>5 (2a)</p>	<p>Denominator Statement: See numerator section for all data needed, and codes</p> <p>Time Window:</p> <p>Denominator Details (Definitions, codes with description):</p>
<p>6 (2a, 2d)</p>	<p>Denominator Exclusions: None</p> <p>Denominator Exclusion Details (Definitions, codes with description):</p>
<p>7 (2a, 2h)</p>	<p>Stratification Do the measure specifications require the results to be stratified? No ► If "other" describe:</p> <p>Identification of stratification variable(s):</p> <p>Stratification Details (Definitions, codes with description):</p>
<p>8 (2a, 2e)</p>	<p>Risk Adjustment Does the measure require risk adjustment to account for differences in patient severity before the onset of care? No ► If yes, (select one) ► Is there a separate proprietary owner of the risk model? No</p> <p>Identify Risk Adjustment Variables: See section 28 for rationale and support for not risk adjusting this measure. Measure was tested against risk adjusted mortality--details on that provided in Section 26.</p>

	Detailed risk model: attached <input type="checkbox"/> OR Web page URL:
9	Type of Score: Rate/proportion Calculation Algorithm: attached <input checked="" type="checkbox"/> OR Web page URL:
(2a)	Interpretation of Score (<i>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</i>) Better quality = Score within a defined interval ► If "Other", please describe:
10	Identify the required data elements (e.g., primary diagnosis, lab values, vital signs): procedure codes, diagnosis codes
(2a, 4a, 4b)	Data dictionary/code table attached <input checked="" type="checkbox"/> OR Web page URL: Data Quality (2a) <i>Check all that apply</i> <input checked="" type="checkbox"/> Data are captured from an authoritative/accurate source (e.g., lab values from laboratory personnel) <input checked="" type="checkbox"/> Data are coded using recognized data standards <input checked="" type="checkbox"/> Method of capturing data electronically fits the workflow of the authoritative source <input type="checkbox"/> Data are available in EHRs <input checked="" type="checkbox"/> Data are auditable
11	Data Source and Data Collection Methods <i>Identifies the data source(s) necessary to implement the measure specifications. Check all that apply</i>
(2a, 4b)	<input type="checkbox"/> Electronic Health/Medical Record <input type="checkbox"/> Electronic Clinical Database, Name: <input type="checkbox"/> Electronic Clinical Registry, Name: <input checked="" type="checkbox"/> Electronic Claims <input type="checkbox"/> Electronic Pharmacy data <input type="checkbox"/> Electronic Lab data <input type="checkbox"/> Electronic source - other, Describe: <input type="checkbox"/> Paper Medical Record <input type="checkbox"/> Standardized clinical instrument, Name: <input type="checkbox"/> Standardized patient survey, Name: <input type="checkbox"/> Standardized clinician survey, Name: <input checked="" type="checkbox"/> Other, Describe: Collected directly from hospitals who utilize administrative claims data to report on 12 month period. Instrument/survey attached <input checked="" type="checkbox"/> OR Web page URL:
12	Sampling <i>If measure is based on a sample, provide instructions and guidance on sample size.</i> Minimum sample size: h1
(2a)	Instructions:
13	Type of Measure: Outcome ► If "Other", please describe:
(2a)	► If part of a composite or paired with another measure, please identify composite or paired measure While the measure to two types of information components, the results are not a composite as is defined by NQF, but rather a reliability adjusted measure of survival. Volume is used to create a volume predicted mortality for the hospital--this component of the measure is used to create greater reliability for low-volume hospitals. In the modeling for this measure, the volume predicted mortality and the observed mortality are weighted. In the model, lower volume hospitals have a higher weight on the volume predicted mortality versus the observed mortality. The opposite is true for high volume hospitals, which have a higher weight on the observed mortality. This methodology results in a reliability adjusted survival predictor.
14	Unit of Measurement/Analysis (<i>Who or what is being measured</i>) <i>Check all that apply.</i>
(2a)	<input type="checkbox"/> Can be measured at all levels <input type="checkbox"/> Individual clinician (e.g., physician, nurse) <input type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice) <input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home) <input type="checkbox"/> Integrated delivery system <input checked="" type="checkbox"/> Health plan <input checked="" type="checkbox"/> Community/Population <input type="checkbox"/> Other (<i>Please describe</i>):
15	Applicable Care Settings <i>Check all that apply</i>
(2a)	<input type="checkbox"/> Can be used in all healthcare settings <input type="checkbox"/> Ambulatory Care (office/clinic) <input type="checkbox"/> Behavioral Healthcare <input type="checkbox"/> Hospice <input checked="" type="checkbox"/> Hospital <input type="checkbox"/> Long term acute care hospital

<input type="checkbox"/> Community Healthcare <input type="checkbox"/> Dialysis Facility <input type="checkbox"/> Emergency Department <input type="checkbox"/> EMS emergency medical services <input type="checkbox"/> Health Plan <input type="checkbox"/> Home Health	<input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF) <input type="checkbox"/> Prescription Drug Plan <input type="checkbox"/> Rehabilitation Facility <input type="checkbox"/> Substance Use Treatment Program/Center <input type="checkbox"/> Other (<i>Please describe</i>):
---	---

IMPORTANCE TO MEASURE AND REPORT

Note: This is a threshold criterion. If a measure is not judged to be sufficiently important to measure and report, it will not be evaluated against the remaining criteria.

16
(1a) Is measure related to a National Priority Partners priority area? [Safety reliability](#) (for NQF staff use) Does measure address a specific NPP goal? (www.qualityforum.org/about/NPP/):

17
(1a) Does the measure address a high impact aspect of healthcare patient/societal consequences of poor quality

Summary of Evidence: This measure addresses mortality in an extremely high risk procedure (pancreatic resection) and is an outcome measure which is of interest to both consumers and purchasers.

While this is a relatively low volume procedure, it is one that has great variation in mortality across hospitals. The absolute difference between low and high volume hospital mortality exceeded 10% (Birkmeyer, et al. [4]. As indicated, mortality in US hospitals varies for pancreatic resection surgeries-- there are significant documented differences between high and low performing hospitals [4]. Higher volumes are associated with better outcomes including lower mortality.

This measure improves upon the technology of surgical mortality measurement. It overcomes three problems with existing mortality measures: 1) Mortality rates are often too "noisy" to reflect hospital quality with surgery (particularly among lower volume hospitals), 2) volume alone is a weak proxy for most procedures, and 3) when both volume and mortality are reported as separate indicators it is difficult to understand which measure is more important. [1] Given the relatively small number of esophagectomy procedures performed annually in the United States, it is important that a mortality measure is designed to reliably measure low volume hospitals, and this measure specifically addresses hospitals which perform relatively few procedures. Up to this point, in order to measure this outcome, other measure developers have added less significant procedures to the denominator in order to gain reliability. Yet, consumers and purchasers would benefit more from knowing specifically where to get one of the most risky procedures performed. The information from the survival predictor is more reliable for these small volume counts than existing measures. In addition, this measure can be applied to the nation, states, or regions.

Birkmeyer and Dimick (2009)[4] show that differences in mortality can be predicted using a reliability adjusted mortality rate (a weighted combination of volume and mortality) which is particularly relevant for selective-referral or public reporting contexts. They reduce the effects of random chance (statistical noise) and as a result with CABG, for example, more than half of the observed variation can be attributed to statistical noise. When they sorted hospitals simply on actual (risk-adjusted) mortality, rates varied from 1.4% to 11.0% across hospital quintiles (Figure 1 in White Paper [1]). After they adjusted for reliability, however, the mortality rates varied considerably less, from 3.3% to 6.3%. Although the almost twofold variation in mortality still suggests ample opportunity for quality improvement, these data underscore the importance of accounting for chance in understanding variation in hospital outcomes.

Citations² for Evidence:

[1] DeFrances, C.J., Lucas, CA, Bule, VC., Golosinskiy, A. 2006 National Hospital Discharge Survey,

² Citations can include, but are not limited to journal articles, reports, web pages (URLs).
 NQF Measure Submission Form, V3.1

	<p>National health statistics reports, no. 5. Hyattsville, MD: National Center for Health Statistics. 2008. Accessed on 12/17/08 at http://www.cdc.gov/nchs/data/nhsr/nhsr005.pdf</p> <p>[2] The National Hospital Bill: The Most Expensive Conditions by Payer, 2006. Statistical Brief #59. File accessed on March 16, 2009, at: http://www.hcup-us.ahrq.gov/reports/statbriefs/sb59.jsp Produced by AHRQ, Center for Delivery, Organization, and Markets, Healthcare Cost and Utilization Project, Nationwide Inpatient Sample, 2006.</p> <p>[3] Composite Measures for Predicting Hospital Mortality with Surgery. Dimick, J.B. Birkmeyer, J.D., White Paper, February 2008, access at: http://www.leapfroggroup.org/media/file/SurvivalPredictorWhitepaper.pdf</p> <p>[4] Birkmeyer, J.D., and Dimick, J.B. (2009) Understanding and reducing variation in surgical mortality. <i>Annu. Rev. Med.</i> 2009. 60:405-15.</p>
<p>18 (1b)</p>	<p>Opportunity for Improvement <i>Provide evidence that demonstrates considerable variation, or overall poor performance, across providers.</i></p> <p>Summary of Evidence: In 2002, a systematic review of the literature on the volume-outcome relationship found that there was a significant relationship between hospital volume and outcomes for pancreatic resection surgery. Unlike the relationship for CABG which was less robust, both surgical procedures for esophagectomy and pancreatectomy were robust. [5]</p> <p>Given the findings related to volume of procedures, Silber et al., [7] explored the relative contribution of complication rates and failure to rescue rates to mortality and found that complication rates were more likely influenced by patient factors while failure to rescue rates of those with complications was more related to hospital factors. Thus, it may be that higher volume hospitals are better at rescuing patients with complications. Silbers finding, in conjunction with the volume information, suggests lower volume hospitals with worse mortality rates could in fact address this through better care following the procedure, thereby reducing their overall rate.</p> <p>Note: Birkmeyer and Dimmick [4] indicate it is also likely that some lower volume hospitals would also have lower mortality rates.</p> <p>Citations for Evidence: [5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. <i>Annals of Internal Medicine</i>, Sept 1;137(6):511-20</p> <p>[6] Epstein, A.J., Rathore, S.S., Krumholz, H.M., and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. <i>BMC Health Services Research</i>, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>[7] Silber, J.H., Rosenbaum, P.R., Trudeau, M.E., et al. 2005. Changes in prognosis after the first postoperative complication. <i>Medical Care</i>, 43:122-31.</p>
<p>19 (1b)</p>	<p>Disparities <i>Provide evidence that demonstrates disparity in care/outcomes related to the measure focus among populations.</i></p> <p>Summary of Evidence: It is more likely that minorities will be treated at a low volume facility, and as a result are likely to be impacted by higher mortality rates. In an analysis of the National Inpatient Sample, Epstein, Rathore and Krumholz (2005)[6] found that a greater proportion of patients treated in low volume hospitals complex cardiovascular procedures were non-white, while a lower proportion of non-white patients presented as "elective" admissions or patients received in transfer as compared to patients in high volume hospitals. It is likely that patients with a complex pancreatic resection would be similarly treated.</p> <p>Citations for evidence:</p>

	<p>[6] Epstein, A.J., Rathore, S.S., Krumholz, H.M., and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. BMC Health Services Research, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p>
<p>20 (1c)</p>	<p>If measuring an Outcome Describe relevance to the national health goal/priority, condition, population, and/or care being addressed: A pancreatic resection is a high risk procedure, and an expensive procedure, and only limited information is available nationally on the risk of mortality associated with the procedure. Other entities with clinical information are not publicly reporting mortality rates of pancreatic resection procedures by hospital provider. This measure is designed to give feedback to hospitals across the country as well as to provide information for decision-making by consumers and purchasers.</p> <p>Mortality in US hospitals varies for pancreatic resection surgeries--there are documented differences between high and low performing hospitals [4,5a]. Higher volumes are associated with better outcomes including lower mortality. In addition to being a high risk surgery, this surgery is one of the high cost procedures.</p> <p>This measure is highly relevant to both consumers and purchasers, given its high cost both in terms of lives lost and dollars spent. National purchasers are interested in comparative information on hospitals nationwide. Pauly (1996) in a study of purchaser interests in hospital performance reporting found that mortality ratings were more important to purchasers than were morbidity or complications. [9] Health plans are interested in contracting with centers of excellence, which can be identified through the results of survival predictor in combination with other information on cost and quality. Consumers have shown their interest in other surgical mortality by requesting reports from the state of Pennsylvania [10]; an earlier study by IOM (Lohr, Donaldson and Walker 1991) found that consumers were interested in hospital mortality rates, but did not perceive this information to be available.[11] Hibbard and Jewett found that consumers were more interested in "undesirable events" (such as mortality, complications, infections) than in "desirable events." [12]</p> <p>[9] Pauly, M.V., Brailer, D.J.Kroch, E., and Even-Shoshan, O. Measuring Hospital Outcomes from a Buyer's Perspective. American Journal of Medical Quality, 11(8): Fall 1996.</p> <p>[10] Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.</p> <p>[11]Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. Quality Review Bulletin 17:242-53.</p> <p>[12]Hibbard, J.H. and Jewett, J.(1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? Medical Care Research and Review., Vol 53(1): 28-47.</p> <p>If not measuring an outcome, provide evidence supporting this measure topic and grade the strength of the evidence <i>Summarize the evidence (including citations to source) supporting the focus of the measure as follows:</i></p> <ul style="list-style-type: none"> • <u>Intermediate outcome</u> - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit. • <u>Process</u> - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s). • <u>Structure</u> - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit. • <u>Patient experience</u> - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public. • <u>Access</u> - evidence that an association exists between access to a health service and the outcomes of,

	<p>or experience with, care.</p> <ul style="list-style-type: none"> • Efficiency- demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality. <p>Type of Evidence <i>Check all that apply</i></p> <table border="0"> <tr> <td><input type="checkbox"/> Evidence-based guideline</td> <td><input checked="" type="checkbox"/> Quantitative research studies</td> </tr> <tr> <td><input checked="" type="checkbox"/> Meta-analysis</td> <td><input type="checkbox"/> Qualitative research studies</td> </tr> <tr> <td><input checked="" type="checkbox"/> Systematic synthesis of research</td> <td><input type="checkbox"/> Other (<i>Please describe</i>):</td> </tr> </table> <p>Overall Grade for Strength of the Evidence³ (<i>Use the USPSTF system, or if different, also describe how it relates to the USPSTF system</i>): Moderate</p> <p>Summary of Evidence (<i>provide guideline information below</i>): Over 100 articles published related to volume and outcome relationship, with some inconsistency in results. Systematic review of the literature conducted in 2002. No review since that time.</p> <p>Citations for Evidence:</p> <p>[5] Halm, EA, Lee C, Chassin, M.R., (2002). Is volume related to outcome in health care? A Systematic Review and methodologic critique of the literature. <i>Annals of Internal Medicine</i>, Sept 1;137(6):511-20.</p> <p>[14] Birkmeyer, J.D., Dimick, J.B., Staiger, D.O. (2006) Operative mortality and procedure volume as predictors of subsequent hospital performance. <i>Ann Surg</i>. 243:411-417.</p> <p>[15] Dimick, JB, Welch HG, Birkmeyer JD. (2004) Surgical mortality as an indicator of hospital quality: The problem with small sample size. <i>JAMA</i>, 292:847-851.</p> <p>[4] Birkmeyer, JD., and Dimick, JB. (2009) Understanding and reducing variation in surgical mortality. <i>Annu. Rev. Med.</i> 60:405-15.</p> <p>[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. <i>Medical Care</i>, 47(2): 226-233.</p> <p>[18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. <i>N Engl J Med</i>. 1979;301:1364-9.</p>	<input type="checkbox"/> Evidence-based guideline	<input checked="" type="checkbox"/> Quantitative research studies	<input checked="" type="checkbox"/> Meta-analysis	<input type="checkbox"/> Qualitative research studies	<input checked="" type="checkbox"/> Systematic synthesis of research	<input type="checkbox"/> Other (<i>Please describe</i>):
<input type="checkbox"/> Evidence-based guideline	<input checked="" type="checkbox"/> Quantitative research studies						
<input checked="" type="checkbox"/> Meta-analysis	<input type="checkbox"/> Qualitative research studies						
<input checked="" type="checkbox"/> Systematic synthesis of research	<input type="checkbox"/> Other (<i>Please describe</i>):						
<p>21 (1c)</p>	<p>Clinical Practice Guideline <i>Cite the guideline reference; quote the specific guideline recommendation related to the measure and the guideline author's assessment of the strength of the evidence; and summarize the rationale for using this guideline over others.</i></p> <p>Guideline Citation:</p> <p>Specific guideline recommendation:</p> <p>Guideline author's rating of strength of evidence (<i>If different from USPSTF, also describe it and how it relates to USPSTF</i>):</p> <p>Rationale for using this guideline over others:</p>						

³The strength of the body of evidence for the specific measure focus should be systematically assessed and rated, e.g., USPSTF grading system www.ahrq.gov/clinic/uspstmeth.htm: **A** - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. **B** - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. **C** - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. **D** - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. **I** - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

<p>22 (1c)</p>	<p>Controversy/Contradictory Evidence Summarize any areas of controversy, contradictory evidence, or contradictory guidelines and provide citations.</p> <p>Summary: There are three areas of possible contention with this measure--</p> <p>1) The volume-outcome relationship has been questioned for some procedures [6, 17]</p> <p>More than 100 studies have demonstrated better results at high-volume hospitals with cardiovascular surgery, major cancer resections (pancreatic resection), and other high-risk procedures.[18, 20]</p> <p>All studies listed here were done to determine whether there was a volume-outcome relationship for hospitals performing surgical procedures--they all documented that there were differences between low volume hospitals and high volume hospitals in mortality--and the evidence for this relationship appears strongest for two procedures, esophagectomy and pancreatectomy with high volume hospitals having less mortality. In the case of esophagectomy the risk of dying was 5-fold more at low volume hospitals. [5a]</p> <p>2) That outcome measures must be risk-adjusted unless there is evidence to show it is not needed (NQF). The survival predictor measure predicts better than volume or mortality alone, and is as good a predictor as risk-adjusted mortality. When testing the unadjusted survival predictor against risk-adjusted mortality there was a (.96) correlation. [4] See Section 28 of this form for details.</p> <p>3) The weighting of input measures into composites. Existing approaches rely on overly simplistic approaches. Among these, assigning equal weight to all measures (i.e., the all or none approach) and relying on expert opinion are the most common. The survival predictor relies on empiric methods for weighting the input measures.</p> <p>Citations:</p> <p>[6] Andrew J Epstein, Saif S Rathore, Harlan M Krumholz and Kevin GM Volpp. Volume-based referral for cardiovascular procedures in the United States: A cross-sectional regression analysis. BMC Health Services Research, 2005, vol. 5:42, accessed at: http://www.biomedcentral.com/1472-6963/5/42</p> <p>[5a] Dimick, J.B., Pronovost, P.J., Cowan, J.A., and Lipsett, P.A. (2003). Surgical volume and quality of care for esophageal resection: do high volume hospitals have fewer complications? Ann Thorac Surg., 75:337-341.</p> <p>[16]Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. Medical Care, 47(2): p. 232.</p> <p>[17] Edward L. Hannan, PhD; Chuntao Wu, PhD; Thomas J. Ryan, MD; Edward Bennett, MD; Alfred T. Culliford, MD; Jeffrey P. Gold, MD; Alan Hartman, MD; O. Wayne Isom, MD; Robert H. Jones, MD; Barbara McNeil, MD, PhD; Eric A. Rose, MD; Valavanur A. Subramanian, MD. Do Hospitals and Surgeons With Higher Coronary Artery Bypass Graft Surgery Volumes Still Have Lower Risk-Adjusted Mortality Rates? Circulation. 2003;108:795-801.</p> <p>[18] Luft HS, Bunker JP, Enthoven AC. Should operations be regionalized? The empirical relation between surgical volume and mortality. N Engl J Med. 1979;301:1364-9.</p> <p>[</p> <p>[20] Begg CB, Cramer LD, Hoskins WJ, Brennan MF. Impact of hospital volume on operative mortality for major cancer surgery. JAMA. 1998;280:1747-51.)</p>
<p>23 (1)</p>	<p>Briefly describe how this measure (as specified) will facilitate significant gains in healthcare quality related to the specific priority goals and quality problems identified above:</p> <p>This measure of predicted survival improves upon the reliability of mortality results for high risk surgical procedures, such as pancreatic resection. For the first time, this measure produces reliable mortality/suivability information on smaller volume hospitals, as well as high volume hospitals. Hospitals across the country will have information available through voluntary public reporting.</p>

SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES	
	<p>Note: Testing and results should be summarized in this form. However, additional detail and reports may be submitted as supplemental information or provided as a web page URL. If a measure has not been tested, it is only potentially eligible for time-limited endorsement.</p>
24	<p>Supplemental Testing Information: attached <input checked="" type="checkbox"/> OR Web page URL:</p>
25	<p>Reliability Testing</p> <p>(2b) Data/sample: Data was a 100% sample from the Medicare Analysis Provider and Review (MEDPAR) files for 2000-2003, these files contain 100% of Medicare hospitalizations for years specified. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the pancreatic resection surgery. We selected only those pancreatic resections where there was an associated cancer diagnosis present, thereby creating a more homogenous risk pool [3].</p> <p>Note: Needleman, Buerhaus, et al., (2003) concluded after applying operational tests on Medicare data for adverse outcomes and all-patient hospital data from 11 states, that Medicare data could be used to assess quality in hospitals.[20] Given the lack of a national all-patient database, MEDPAR data was used in development and testing of the models.</p> <p>Analytic Method:</p> <p>Model Development</p> <p>We used an empirical Bayes approach to combine mortality rates with information on hospital volume at each hospital. In traditional empirical Bayes methods, a point estimate (e.g., mortality rate observed at a hospital) is adjusted for reliability by shrinking it towards the overall mean (e.g., overall mortality rate in the population) [21,22]. We modified this traditional approach by shrinking the observed mortality rate back toward the mortality rate expected given the volume at that hospital—we refer to this as the “volume-predicted mortality” (See attached White Paper TECHNICAL APPENDIX for the mathematical details of this method). With this approach, the observed mortality rate is weighted according to how reliably it is estimated, with the remaining weight placed on the information regarding hospital volume. Because this method includes observed data to the extent that it is useful, and only relies on the proxy measure to the extent necessary, it ensures an optimal combination of these two quality domains. [3]</p> <p>The two inputs to the survival predictor measure are mortality rates and procedure volume for each of the six included operations. Procedure-specific mortality rates were calculated for all hospitals over a 2-year period (2000-01) and this was used as the first input. Hospital volume was calculated as the number of Medicare cases performed during the same time period. For each operation, the relationship between hospital volume and risk-adjusted mortality was modeled using linear regression. (Details of the risk-adjustment strategy will be discussed below.) After testing the fit of several transformations, hospital volume was modeled as the natural log of the continuous volume variable, which is the same approach used in our previous work [23]. Using this regression model, we estimated the volume-predicted mortality, the second input to the survival predictor measure.</p> <p>We then used the empirical Bayes approach to create an optimal combination of these two inputs. This survival predictor measure theoretically provides the best estimate of a hospitals true mortality rate, taking into account the both available inputs [21,22]. The combined survival predictor measure was calculated as follows: mortality prediction = (weight)*(observed mortality) + (1-weight)*(volume-predicted mortality). The weight placed on the point estimate of mortality is the reliability, or ratio of signal to</p>

	<p>signal plus noise, calculated as follows: weight = variation among hospitals/(variation among hospitals + variation within hospitals). The variation among hospitals was calculated as the variance in observed mortality rates for the hospitals included in the sample. The variation within hospitals was calculated as the standard error of the mortality rate at each hospital. With this method, more weight is placed on the observed mortality rate when a hospital has a high number of cases because it is estimated with more reliability; less weight is placed on the observed mortality rate when a hospital performs a low number of cases because of its lower reliability.</p> <p>A calculation worksheet is attached.</p> <p>Testing Results: Hospital caseloads and the weights applied to each input to the survival predictor measure varied for each procedure studied (see Table 1 in white paper [3]). For pancreatic resection surgical procedures, a procedure with relatively low hospital caseloads, the weight applied to the volume input was .77 ([3]-Table 1), indicating that the observed mortality was less reliable than the volume predicted mortality for hospitals with low volumes.</p> <p>The survival predictor (mortality) measure explained a large proportion of non-random, hospital-level variation in risk-adjusted mortality rates (see Table 2 in White Paper [3]). For pancreatic resection procedures, the survival predictor explained 59% of the hospital level variation in mortality rates; this compares to 23% for observed mortality and 57% for volume of pancreatic resection surgeries. Measures with low reliability or correlation explain little variation. The correlation between the survival predictor and risk-adjusted mortality was (.96) ([16] p. 232), and the amount of variation explained was 59% [3]. This is an adequate level of reliability.</p> <p>Note: The percentage of hospital level variation explained by the predictor is equivalent to an R-squared from a regression. ([16]-p. 228)</p> <p>Citations:</p> <p>[3] p. 19 (Table 2)</p> <p>[16] Staiger, D., Dimick, J., Baser, O., Fan, Z., and Birkmeyer, J. 2009. Empirically Derived Composite Measures of Surgical Performance. <i>Medical Care</i>, 47(2): 226-233</p> <p>[21] Morris CN. Parametric Empirical Bayes Inference: Theory and Applications. <i>J Am Stat Assoc</i> 1988;78:47-55.</p> <p>[22] McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) <i>Frontiers in Health Policy Research</i>. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.</p> <p>[23] Birkmeyer JD, Stukel TA, Siewers AE, et al. Surgeon volume and operative mortality in the United States. <i>N Engl J Med</i>. 2003;349:2117-2127.</p> <p>[20] Needleman, J., Buerhaus, P.I., Mattke, S., Stewart, M., and Zelevinsky, M. (2003). <i>Health Services Research</i> 38.6, Part I; 1487-1508.</p>
<p>26</p> <p>(2c)</p>	<p>Validity Testing</p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p>

Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the pancreatic resection surgery. We only included patients with cancer as a diagnosis (which is the bulk of the population and highly correlated with the full population).

Analytic Method: We determined the value of our survival predictor (mortality) measure by establishing whether it explained hospital-level variation in risk-adjusted mortality rates and by assessing to what degree it was able to predict future hospital performance. We first estimated the proportion of variation in hospital-level mortality (2000-01) explained by the survival predictor measure using random effects logistic regression models. For these analyses, we estimated the proportional change in the hospital-level variance in mortality rates, which was determined from the standard deviation of the random effect, after adding each measure to the model [14,22].

We next compared the ability of the survival predictor measure to the individual measures, mortality rates and hospital volume. We should note that these analyses focus on explaining systematic, or non-random, variation, since measurement error (random error) is accounted for and subtracted from the total variation in all analyses [22,24].

We next determined the extent to which the composite measure predicts future risk-adjusted mortality. For this analysis, hospitals were ranked based on each measure from the earlier time period (data from years 2000-01) and divided into four equal size groups (quartiles at the patient level). The subsequent risk-adjusted mortality rates for each quartile of performance were then calculated (data from years 2002-03). We present the subsequent mortality rates across quartiles of the pancreatic resection survival predictor measure to graphically demonstrate its usefulness in discriminating among hospitals for the entire spectrum of performance. To compare the predictive ability of the composite measures and individual measures, we also present the subsequent mortality rates in the "worst" compared to the "best" quartile in the White Paper ([3], p. 22) "Quartiles of Performance Measures (2000-2001). This table reflects how well the unadjusted survival predictor created on 2000-2001 data compares to risk-adjusted mortality in 2002-2003 data.

Note: The risk-adjusted mortality rate for esophagectomy was constructed using standard methods. We determined the ratio of actual deaths or complications to the number of expected deaths (the O/E ratio). The number of expected deaths was the sum over all patients of the predicted probability of death or complications derived from a logistic regression model estimated on all patients undergoing pancreatic resection surgery. The dependent variable in the logistic model was death or complications and the independent variables were patient covariates. The patient characteristics included age, gender, race, admission acuity, and co-existing diseases using the Elixhauser method. A zip code level measure of socio-economic status was derived from 2000 census data.

Testing Results: While some measures are good at discriminating top performers or bottom performers, this measure is good at prediction across entire spectrum of performance. [See White paper [3]: Figures p. 21-22) for a graphical demonstration of the usefulness of the survival predictor in discriminating among hospitals across the entire spectrum of performance.] To compare the predictive ability of the reliability adjusted survival predictor versus the individual components (volume and observed mortality) we also present the subsequent mortality rates in the "worst" compared to the "best" quartile.

[22]. McClellan MB, Staiger DO. Comparing the Quality of Health Care Providers. Alan Garber (ed.) Frontiers in Health Policy Research. Volume 3. 2000 The MIT Press: Cambridge MA, pp. 113-136.

[14] Birkmeyer JD, Dimick JB, Staiger DO. Operative mortality and procedure volume as predictors of subsequent hospital performance. *Ann Surg* 2006;243:411-417.

[24] Zaslavsky AM, Cleary PD. Dimensions of plan performance for sick and healthy members on the Consumer Assessments of Health Plans Study 2.0 survey. *Med Care* 2002;40:951-964.

<p>27 (2d)</p>	<p>Measure Exclusions <i>Provide evidence to justify exclusion(s) and analysis of impact on measure results during testing.</i></p> <p>Summary of Evidence supporting exclusion(s): The developers defined the denominator to minimize potential for case mix differences between hospitals, they created homogenous sub-groups, in this case only those pancreatic resections with a cancer diagnosis. This essentially left a relatively homogenous population undergoing elective, non-emergency surgeries.</p> <p>Only those hospitals with elective cases will have a survival predictor, since the primary goal of the measure is to provide information for selection of a specific hospital for the pancreatic resection procedure.</p> <p>Citations for Evidence:</p> <p>Data/sample:</p> <p>Analytic Method:</p> <p>Testing Results:</p>
<p>28 (2e)</p>	<p>Risk Adjustment Testing <i>Summarize the testing used to determine the need (or no need) for risk adjustment and the statistical performance of the risk adjustment method.</i></p> <p>Data/sample: Data from the Medicare Analysis Provider and Review (MEDPAR) files, which contains 100% of Medicare hospitalizations. MEDPAR files, which contain hospital discharge abstracts for all fee-for-service acute care hospitalizations of all US Medicare recipients, were used to create our main analysis datasets. The Medicare eligibility file was used to assess patient vital status at 30 days. The study protocol was approved by the Institutional Review Board at the University of Michigan.</p> <p>Using appropriate procedure codes from the International Classification of Diseases, version 9 (ICD-9), we identified all patients aged 65 to 99 undergoing the pancreatic resection surgery. We created homogenous patient subgroups, including those with diagnosis codes indicating that the patient had a related cancer.</p> <p>Analytic Method: Sensitivity analysis. We performed a sensitivity analysis to determine whether risk-adjustment of the mortality input was important in improving the predictive ability of the survival predictor measure. Risk-adjustment was performed using logistic regression to estimate expected mortality rates for each hospital based on patient age, gender, race, urgency of operation, median income, and coexisting diseases. Coexisting diseases were determined from secondary diagnostic codes using the methods of Elixhauser (16). The observed mortality rate at each hospital was then divided by the expected mortality rate to yield the ratio of observed/expected deaths (O/E ratio). The O/E ratio was multiplied by the average mortality rate for each operation to yield a risk-adjusted mortality rate. To determine the value of risk-adjustment in the context of selective referral, we compared the ability of risk-adjusted and unadjusted composite measures to predict subsequent performance.</p> <p>Testing Results: In sensitivity analysis, composite measures based on an unadjusted mortality input and a risk-adjusted mortality input had a correlation of (.95) and thus were equally good at predicting future performance (See pages 21-22 in the White Paper [3]).</p> <p>► If outcome or resource use measure not risk adjusted, provide rationale: Because risk-adjusted mortality is not available publicly except for limited locations, the capacity to use unadjusted mortality is very desirable, especially since it was shown to provide (under this methodology) an equal result. This measure will allow measurement to occur across the United States, providing information to national companies, health plans and consumers.</p>
<p>29 (2g)</p>	<p>Testing comparability of results when more than 1 data method is specified (e.g., administrative claims or chart abstraction)</p> <p>Data/sample: not applicable</p>

	<p>Analytic Method:</p> <p>Results:</p>
30 (2f)	<p>Provide Measure Results from Testing or Current Use Results from testing</p> <p>Data/sample: same as described above, results for survival predictor in White Paper [3]available on Website and Validation results for composite in [16] Staiger, Dimick et al., Medical Care 2009</p> <p>Methods to identify statistically significant and practically/meaningfully differences in performance: Bayesian Hierarchical methods using new shrinkage estimator Empirical Bayesian methods to determine weights Correlations Calculated the amount of variation predicted by survival predictor as a percentage of all hospital-level variation (adjusted for sampling variation)--analogous to a R-squared from a regression that summarizes the ability of the predictor to explain the hospital level variation in mortality for pancreatic resection surgery. Predictor was tested against the "gold standard" --risk adjusted mortality</p> <p>Results: See White Paper [3]</p>
31 (2h)	<p>Identification of Disparities</p> <p>► If measure is stratified by factors related to disparities (i.e. race/ethnicity, primary language, gender, SES, health literacy), provide stratified results:</p> <p>► If disparities have been reported/identified, but measure is not specified to detect disparities, provide rationale: .</p>
USABILITY	
32 (3)	<p>Current Use Testing completed If in use, how widely used Nationally ► If "other," please describe: Survival Predictor for Pancreatectomy and Esophagectomy in use--see URL.</p> <p><input checked="" type="checkbox"/> Used in a public reporting initiative, name of initiative: Leapfrog Hospital Survey Sample report attached <input type="checkbox"/> OR Web page URL: https://www.leapfroggroup.org/cp</p>
33 (3a)	<p>Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)</p> <p>Data/sample:</p> <p>Methods:</p> <p>Results: See following citations reflecting consumer use of mortality information:</p> <p>[10]Pennsylvania Health Care Cost Containment Council. (1993). A progress report 1991-1993: The use of the council's information and its impact on the cost and quality of healthcare. Harrisburg, PA.</p> <p>[11]Lohr, K., Donaldson, M., and Walker, A. (1991). Medicare: A strategy for quality assurance, III: Beneficiary and physician focus groups. Quality Review Bulletin 17:242-53.</p> <p>[12]Hibbard, J.H. and Jewett, J.(1996). What Type of Quality Information Do Consumers Want in a Health Care Report Card? Medical Care Research and Review., Vol 53(1): 28-47.</p>
34 (3b, 3c)	<p>Relation to other NQF-endorsed™ measures</p> <p>► Is this measure similar or related to measure(s) already endorsed by NQF (on the same topic or the same target population)? <i>Measures can be found at www.qualityforum.org under Core Documents.</i></p> <p><i>Check all that apply</i></p> <p><input type="checkbox"/> Have not looked at other NQF measures <input checked="" type="checkbox"/> Other measure(s) on same topic <input checked="" type="checkbox"/> Other measure(s) for same target population <input type="checkbox"/> No similar or related measures</p>

	<p>Name and number of similar or related NQF-endorsed™ measure(s): Pancreatic Resection Volume IQI2 NQF #0366; Pancreatic Resection Mortality Rate (IQI9) NQF # 0365.</p> <p>Are the measure specifications harmonized with existing NQF-endorsed™ measures? Not harmonized ▶ If not fully harmonized, provide rationale: This new measure requires a diagnosis of cancer in addition to the procedure codes--the cohort of patients with cancer can electively select site for surgery. If performed as an emergent case, patients do not elect where to have the procedure. We are concerned about elective procedures where patients have no information on where else they could go for better outcomes.</p> <p>Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures: This measure provides the ability to produce reliable mortality results for low volume hospitals, other measures do not have this capacity. In addition, the access to data nationally for other pancreatic resection mortality measures does not exist in the public domain.</p>
FEASIBILITY	
35 (4a)	<p>How are the required data elements generated? <i>Check all that apply</i></p> <p><input checked="" type="checkbox"/> Data elements are generated concurrent with and as a byproduct of care processes during care delivery (e.g., blood pressure or other assessment recorded by personnel conducting the assessment)</p> <p><input type="checkbox"/> Data elements are generated from a patient survey (e.g., CAHPS)</p> <p><input checked="" type="checkbox"/> Data elements are generated through coding performed by someone other than the person who obtained the original information (e.g., DRG or ICD-9 coding on claims)</p> <p><input checked="" type="checkbox"/> Other, Please describe: Data are currently submitted to Leapfrog via a secure online survey-</p>
36 (4b)	<p>Electronic Sources All data elements</p> <p>▶ If all data elements are not in electronic sources, specify the near-term path to electronic collection by most providers:</p> <p>▶ Specify the data elements for the electronic health record: volume of CABG procedure, observed death during inpatient stay, related to CABG procedure</p>
37 (4c)	<p>Do the specified exclusions require additional data sources beyond what is required for the other specifications? No</p> <p>▶ If yes, provide justification:</p>
38 (4d)	<p>Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure: It is unlikely that this procedure, or inpatient death will be inaccurately coded or not coded given the high cost of procedure and the accompanying death.</p> <p>Describe how could these potential problems be audited: If problems were identified, a chart review of cases could be performed.</p> <p>Did you audit for these potential problems during testing? No If yes, provide results:</p>
39 (4e)	<p>Testing feasibility Describe what have you learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues: Initial results only available for Esophagectomy, Pancreatectomy. CABG will be released in 2009</p>
CONTACT INFORMATION	
40	<p>Web Page URL for Measure Information Describe where users (implementers) should go for more details on specifications of measures, or assistance in implementing the measure. Web page URL: https://leapfrog.medstat.com for access to Survival Predictor White Paper</p>
41	<p>Measure Steward Point of Contact First Name: MI: Last Name: Credentials (MD, MPH, etc.):</p>

	<p>Organization: The Leapfrog Group % The Academy Street Address: 1150 17th St., NW, Suite 600 City: Washington State: DC ZIP: 20036 Email: Telephone: ext:</p>
42	<p>Measure Developer Point of Contact If different from Measure Steward First Name: Justin MI: B Last Name: Dimick Credentials (MD, MPH, etc.): MD, MPH Organization: Department of Surgery, University of Michigan, M-SCORE offices, Suite 201 and 202 Street Address: 211 N. Fourth Avenue City: Ann Arbor State: MI ZIP: 48104 Email: jdimick@umich.edu Telephone: ext:</p>
<p>ADDITIONAL INFORMATION</p>	
43	<p>Workgroup/Expert Panel involved in measure development Workgroup/panel used ► If workgroup used, describe the members' role in measure development: Research team led by Justin Dimick, MD, MPH; ► Provide a list of workgroup/panel members' names and organizations:</p> <p>Douglas Staiger Ph.D., Department of Economics and the Dartmouth Institute for Health Policy and Clinical Practice, Dartmouth College, Hanover, New Hampshire</p> <p>John D. Birkmeyer, MD Michigan Surgical Collaborative for Outcomes Research and Evaluation Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Onur Baser, Ph.D. Michigan Surgical Collaborative for Outcomes Research and Evaluation Department of Surgery University of Michigan Ann Arbor, Michigan</p> <p>Research supported by the National Institute on Aging</p>
44	<p>Measure Developer/Steward Updates and Ongoing Maintenance Year the measure was first released: 2008 Month and Year of most recent revision: August 2008 What is the frequency for review/update of this measure? Annual When is the next scheduled review/update for this measure? New coefficients for August 2009</p>
45	<p>Copyright statement/disclaimers: none</p>
46	<p>Additional Information: All measure information is available at https://leapfrog.medstat.com Please contact measure developer prior to use to assure all necessary items have been accessed.</p>
47	<p>I have checked that the submission is complete and any blank fields indicate that no information is provided. <input checked="" type="checkbox"/></p>
48	<p>Date of Submission (MM/DD/YY): Revised submission dated 3/18/09</p>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.0

August 2008

The measure information you submit will be shared with NQF’s Steering Committees and Technical Advisory Panels to evaluate measures against the NQF criteria of importance to measure and report, scientific acceptability of measure properties, usability, and feasibility. Four conditions (as indicated below) must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards. Not all acceptable measures will be strong—or equally strong—among each set of criteria. The assessment of each criterion is a matter of degree; however, all measures must be judged to have met the first criterion, importance to measure and report, in order to be evaluated against the remaining criteria. References to the specific measure evaluation criteria are provided in parentheses following the item numbers. Please refer to the *Measure Evaluation Criteria* for more information at www.qualityforum.org under Core Documents. Additional guidance is being developed and when available will be posted on the NQF website.

Use the tab or arrow (↓→) keys to move the cursor to the next field (or back ←↑). There are three types of response fields:

- drop-down menus - select one response;
- check boxes - check as many as apply; and
- text fields - you can copy and paste text into these fields or enter text; these fields are not limited in size, but in most cases, we ask that you summarize the requested information.

Please note that URL hyperlinks do not work in the form; you will need to type them into your web browser.

Be sure to answer all questions. Fields that are left blank will be interpreted as no or none. Information must be provided in this form. Attachments are not allowed except when specifically requested or to provide additional detail or source documents for information that is summarized in this form. If you have important information that is not addressed by the questions, they can be entered into item #48 near the end of the form.

For questions about this form, please contact the NQF Project Director listed in the corresponding call for measures.

CONDITIONS FOR CONSIDERATION BY NQF	
	<i>Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards.</i>
A (A)	<i>Public domain or Intellectual Property Agreement signed: IP Agreement signed and submitted (If no, do not submit) Template for the Intellectual Property Agreement is available at www.qualityforum.org under Core Documents.</i>
B (B)	<i>Measure steward/maintenance: Is there an identified responsible entity and process to maintain and update the measure on a schedule commensurate with clinical innovation, but at least every 3 years? Yes, information provided in contact section (If no, do not submit)</i>
C (C)	<i>Intended use: Does the intended use of the measure include BOTH public reporting AND quality improvement? Yes (If no, do not submit)</i>
D (D)	<i>Fully developed and tested: Is the measure fully developed AND tested? Yes, fully developed and tested (If not tested and no plans for testing within 24 months, do not submit)</i>

THE NATIONAL QUALITY FORUM

MEASURE SUBMISSION FORM VERSION 3.0

August 2008

	<i>(for NQF staff use)</i> NQF Review #: HOE-004 NQF Project: Hospital Outcomes and Efficiency
MEASURE SPECIFICATIONS & DESCRIPTIVE INFORMATION	
1	<i>Information current as of (date- MM/DD/YY): 10/26/08 original - 3/19/09 updated</i>
2	<i>Title of Measure: RISK-ADJUSTED 30-DAY READMISSION RATE FOR HEART FAILURE</i>
3	Brief description of measure ¹: <i>Assesses the risk-adjusted 30-day readmission rates for patients discharged with heart failure during the measurement year.</i>
4 (2a)	<p>Numerator Statement: <i>Members readmitted to the hospital 2-30 days after the index date.</i></p> <p><i>Note: Index date is defined as the date of discharge from an inpatient setting with congestive heart failure (date of denominator criteria A)</i></p> <p>Time Window: <i>2-3 days after the date of discharge</i></p> <p>Numerator Details (Definitions, codes with description): <i>Numerator Logic : A only</i></p> <p><i>[A] Members readmitted 2-30 days after the index date</i></p> <p>Inpatient setting: <i>CPT-4 codes: 99221-99223, 99261-99263, 99291-99300</i></p>
5 (2a)	<p>Denominator Statement: <i>Members who were discharged from an inpatient setting with congestive heart failure during the 1 year period ending 30 days prior to the end of the measurement year.</i></p> <p>Time Window: <i>One year period ending 30 days prior to the end of the measurement year.</i></p> <p>Denominator Details (Definitions, codes with description): <i>Denominator Logic: DEMO and CE and A</i></p> <p><i>[DEMO] Members ages 19 years and older by end of the measurement year</i></p> <p><i>[CE] Members who are continuously enrolled during the 365 days prior to the index date through 30 days after index date.</i></p> <p><i>Note: Index date is defined as the date of discharge from an inpatient setting with congestive heart failure (date of denominator criteria A)</i></p> <p><i>[A] Members who were discharged from an inpatient setting with congestive heart failure (CHF) during the 1 year period ending 30 days prior to the end of the measurement year.</i></p> <p>CHF: <i>ICD-9 diagnosis codes: 398.91, 402.x1, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, 428.xx</i> AND Inpatient setting: <i>CPT-4 codes: 99238-99239</i></p>

¹ Example of measure description: Percentage of adult patients with diabetes aged 18-75 years receiving one or more A1c test(s) per year.
NQF Measure Submission Form, V3.0

<p>6 (2a, 2d)</p>	<p>Denominator Exclusions: Members in hospice during the 356 days prior to the index date through 30 days after the index date, or members who expired.</p> <p>Note: Index date is defined as the date of discharge from an inpatient setting with congestive heart failure (date of denominator criteria A)</p> <p>Denominator Exclusion Details (Definitions, codes with description): Denominator Exclusion Logic: A or B</p> <p>[A] Members in hospice during the 365 days prior to the index date through 30 days after index date.</p> <p>Note: Index date is defined as the date of discharge from an inpatient setting with congestive heart failure (date of denominator criteria A)</p> <p>ICD-9 diagnosis code: V66.7 CPT-4 codes: 99376*, 99377, 99378 HCPCS codes: G0065*, G0182, G0337, Q5001-Q5009, S0271, S9126, T2042-T2046 UB revenue codes: 115, 125, 135, 145, 155, 235, 650-652, 655-659, 0115, 0125, 0135, 0145, 0155, 0235, 0650-0652, 0655-0659 UB type of bill codes: 81x, 82x Place of service code: 34</p> <p>* Codes is retired but appropriate for retrospective analysis.</p> <p>[B] Members whose discharge status is 'expired' on the index date</p> <p>Note: Index date is defined as the date of discharge from an inpatient setting with congestive heart failure (date of denominator criteria A)</p> <p>Note: codes for discharge status "expired" will vary by plan.</p>
<p>7 (2a, 2h)</p>	<p>Stratification Do the measure specifications require the results to be stratified? No ▶ If "other" describe:</p> <p>Identification of stratification variable(s):</p> <p>Stratification Details (Definitions, codes with description):</p>
<p>8 (2a, 2e)</p>	<p>Risk Adjustment Does the measure require risk adjustment to account for differences in patient severity before the onset of care? Yes ▶ If yes, Statistical Risk Model, see Variables ▶ Is there a separate proprietary owner of the risk model? No</p> <p>Identify Risk Adjustment Variables:</p> <p>(1) Age, Age-squared</p> <p>(2) gender (male vs female)</p> <p>(3) History of congestive heart failure hospitalization in the past year (1 year prior to index date, excluding index date) ICD-9 diagnosis code(s): 398.91, 402.x1, 404.01, 404.03, 404.11, 404.13, 404.91, 404.93, 428.xx AND Inpatient setting: CPT-4 code(s): 99238-99239</p> <p>(4) chronic renal disease (i.e., stage >= 3 or dialyses) (1 year prior to index date, including index date) ICD-9 diagnosis codes: 250.4x, 274.1x, 403.01, 403.11, 403.90, 403.91, 404.02, 404.03, 404.10, 404.11, 404.12, 404.13, 404.90, 404.91, 404.92, 404.93, 581.xx, 582.xx, 583.xx, 585.3-585.5, 586, 587, 753.0, 753.10, 753.11, 753.12, 753.13, 753.14, 753.15, 753.16, 753.17, 753.19</p>

<p>DRG code(s): 316</p> <p>Dialysis: ICD-9 diagnosis codes: 38.95, 39.27, 39.42, 39.93, 39.95, 54.98, V45.1, V56.0, V56.1, V56.2, V56.31, V56.32, V56.8, E879.1</p> <p>ICD-9 surgical proc codes: 38.95, 39.27, 39.42, 39.93, 39.95, 54.98</p> <p>DRG code: 317</p> <p>CPT codes: 0505F, 0507F, 3066F, 3082F-3084F, 4051F-4055F, 36800, 36810, 36815, 36818-36821, 36825, 36831-36833, 90920, 90921, 90924, 90925, 90935, 90937, 90939, 90940, 90945, 90947, 90989, 90993, 90997, 90999, 99512, G0257, G0314-G0319, G0322, G0323, G0326, G0327, G9013, G9014 UB revenue code(s): 0800-0809, 0820-0859, 0880, 0881, 0882, 0889</p> <p>HCPCS codes: A4653, A4671-A4918, E1500-E1699</p> <p>(5) coronary artery disease (1 year prior to index date, including index date) AMI: ICD-9 diagnosis code: 410.x1 DRG codes: 121, 122, 516</p> <p>PTCA: ICD-9 surgical proc codes: 00.66, 36.01, 36.02, 36.05, 36.06, 36.07, 36.09 CPT-4 codes: 33140, 92980-92982, 92984, 92995, 92996 DRG codes: 516, 517, 526, 527, 555-558</p> <p>CABG ICD-9 surgical proc codes: 36.1x, 36.2x HCPCS codes: S2205-S2209 CPT-4 codes: 33510-33514, 33516-33519, 33521-33523, 33533-33536, 35600, 33572 DRG codes: 106, 107, 109, 547-550</p> <p>Other forms of Ischemic Heart Disease: ICD-9 diagnosis codes: 414.0x, 414.8x, 414.9x , 429.2</p> <p>Stable Angina: ICD-9 diagnosis codes: 411.xx, 413.x</p> <p>(6) pacemaker insertion (1 year prior to index date, including index date) CPT codes: 00530, 33226-33240, 33249</p> <p>(7) COPD (1 year prior to index date, including index date) ICD-9 diagnosis codes: 492.x, 506.4, 518.1, 518.2</p> <p>(8) discharge to nursing home (1-30 days after index date) CPT-4 codes: 99304-99316, 99318, 99324-99340</p> <p>(9) Modified Elixhauser Comorbidity Index (codes down loaded from website below; used all claims from 0-365 days prior to index date)that exlcude congestive heart failure, chronic pulmonary disease, and renal failure). http://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp#download</p> <p>Model Description: We used claims data form all available data sources (i.e., facility, professional, demographic, pharmacy, etc.) to construct all covariates. Code details and time frames of measurement are listed above. Because of natural clustering of the observations within hospitals, we used hierarchical generalized linear models (logit link function). We modeled the probabily of readmission within 30 days as a function of patient</p>

	<p>demographic and clinical characteristics and a random hospital specific effect (randome intercept model).</p> <p>Detailed risk model: attached <input type="checkbox"/> OR Web page URL:</p>
9	<p>Type of Score: Rate/proportion Calculation Algorithm: attached <input type="checkbox"/> OR Web page URL:</p> <p>(2a) Interpretation of Score <i>(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)</i> Better quality = Lower score ► If "Other", please describe:</p>
10	<p>Identify the required data elements(e.g., primary diagnosis, lab values, vital signs): Data dictionary/code table attached <input type="checkbox"/> OR Web page URL:</p> <p>(2a. Data Quality (2a) <i>Check all that apply</i></p> <p>4a, 4b) <input type="checkbox"/> Data are captured from an authoritative/accurate source (e.g., lab values from laboratory personnel) <input checked="" type="checkbox"/> Data are coded using recognized data standards <input type="checkbox"/> Method of capturing data electronically fits the workflow of the authoritative source <input type="checkbox"/> Data are available in EHRs <input type="checkbox"/> Data are auditable</p>
11	<p>Data Source and Data Collection Methods <i>Identifies the data source(s) necessary to implement the measure specifications. Check all that apply</i></p> <p>(2a, 4b) <input type="checkbox"/> Electronic Health/Medical Record <input type="checkbox"/> Paper Medical Record <input type="checkbox"/> Electronic Clinical Database, Name: <input type="checkbox"/> Standardized clinical instrument, Name: <input type="checkbox"/> Electronic Clinical Registry, Name: <input type="checkbox"/> Standardized patient survey, Name: <input checked="" type="checkbox"/> Electronic Claims <input type="checkbox"/> Standardized clinician survey, Name: <input type="checkbox"/> Electronic Pharmacy data <input checked="" type="checkbox"/> Other, Describe: Member demographics and member enrollment data <input type="checkbox"/> Electronic Lab data <input type="checkbox"/> Electronic source - other, Describe:</p> <p style="text-align: right;">Instrument/survey attached <input type="checkbox"/> OR Web page URL:</p>
12	<p>Sampling <i>If measure is based on a sample, provide instructions and guidance on sample size.</i> Minimum sample size:</p> <p>(2a) Instructions:</p>
13	<p>Type of Measure: Outcome ► If "Other", please describe:</p> <p>(2a) ► If part of a composite or paired with another measure, please identify composite or paired measure</p>
14	<p>Unit of Measurement/Analysis <i>(Who or what is being measured) Check all that apply.</i></p> <p>(2a) <input type="checkbox"/> Can be measured at all levels <input checked="" type="checkbox"/> Integrated delivery system <input checked="" type="checkbox"/> Individual clinician (e.g., physician, nurse) <input checked="" type="checkbox"/> Health plan <input checked="" type="checkbox"/> Group of clinicians (e.g., facility department/unit, group practice) <input type="checkbox"/> Community/Population <input checked="" type="checkbox"/> Facility (e.g., hospital, nursing home) <input type="checkbox"/> Other <i>(Please describe):</i></p>
15	<p>Applicable Care Settings <i>Check all that apply</i></p> <p>(2a) <input type="checkbox"/> Can be used in all healthcare settings <input type="checkbox"/> Hospice <input type="checkbox"/> Ambulatory Care (office/clinic) <input checked="" type="checkbox"/> Hospital <input type="checkbox"/> Behavioral Healthcare <input type="checkbox"/> Long term acute care hospital <input type="checkbox"/> Community Healthcare <input type="checkbox"/> Nursing home/ Skilled Nursing Facility (SNF) <input type="checkbox"/> Dialysis Facility <input type="checkbox"/> Prescription Drug Plan <input type="checkbox"/> Emergency Department <input type="checkbox"/> Rehabilitation Facility <input type="checkbox"/> EMS emergency medical services <input type="checkbox"/> Substance Use Treatment Program/Center <input checked="" type="checkbox"/> Health Plan <input type="checkbox"/> Other <i>(Please describe):</i> <input type="checkbox"/> Home Health</p>
IMPORTANCE TO MEASURE AND REPORT	
Note: This is a threshold criterion. If a measure is not judged to be sufficiently important to measure	

	and report, it will not be evaluated against the remaining criteria.
16 (1a)	Addresses a Specific National Priority Partners Goal <i>Enter the numbers of the specific goals related to this measure (see list of goals on last page):</i>
17 (1a)	<p>If not related to NPP goal, identify high impact aspect of healthcare patient/societal consequences of poor quality</p> <p>Summary of Evidence:</p> <ul style="list-style-type: none"> Approximately 5 million Americans have heart failure, with over 550,000 new cases diagnosed each year. Heart failure is the primary cause of death for nearly 53,000 patients annually.[1, 2] Heart failure is the most common primary discharge diagnosis among patients older than 65 years, accounting for 1 million hospitalizations and discharges each year.[3-5] 2006 estimates place the combined direct and indirect cost of heart failure in the United States at \$29.6 billion for that year.[6] Among Medicare beneficiaries, the cost of readmissions within 6 months of initial hospitalization averages \$7000 per readmission.[7] <p>Citations² for Evidence:</p> <ol style="list-style-type: none"> Bonow, R.O., et al., ACC/AHA Clinical Performance Measures for Adults with Chronic Heart Failure: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures (Writing Committee to Develop Heart Failure Clinical Performance Measures): endorsed by the Heart Failure Society of America. <i>Circulation</i>, 2005. 112(12): p. 1853-87. Hunt, S.A., et al., 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): developed in collaboration with the American College of Chest Physicians and the International Society for Heart and Lung Transplantation: endorsed by the Heart Rhythm Society. <i>Circulation</i>, 2005. 112(12): p. e154-235. Fonarow, G.C., C.W. Yancy, and J.T. Heywood, Adherence to heart failure quality-of-care indicators in US hospitals: analysis of the ADHERE Registry. <i>Arch Intern Med</i>, 2005. 165(13): p. 1469-77. DeFrances, C.J. and M.J. Hall, 2005 National Hospital Discharge Survey. <i>Adv Data</i>, 2007(385): p. 1-19. Gheorghide, M., et al., Systolic blood pressure at admission, clinical characteristics, and outcomes in patients hospitalized with acute heart failure. <i>Jama</i>, 2006. 296(18): p. 2217-26. Fonarow, G.C., et al., Association between performance measures and clinical outcomes for patients hospitalized with heart failure. <i>Jama</i>, 2007. 297(1): p. 61-70. Phillips, C.O., et al., Comprehensive discharge planning with postdischarge support for older patients with congestive heart failure: a meta-analysis. <i>Jama</i>, 2004. 291(11): p. 1358-67.
18 (1b)	<p>Opportunity for Improvement <i>Provide evidence that demonstrates considerable variation, or overall poor performance, across providers.</i></p> <p>Summary of Evidence:</p> <ul style="list-style-type: none"> The costs associated with treating heart failure, including readmissions, are expected to rise because of the increased aging of the American population.[1] Research shows an underutilization of processes of care such as the use of angiotensin-converting enzyme inhibitors (ACEIs). Lower quality of inpatient care is associated with higher readmission rates and mortality.[2] Readmissions are costly for hospitals and patients. Length-of-stay data suggests readmitted patients account for a disproportionate amount of total ICU costs. In addition, hospital death rates are 2-10 times higher for readmitted patients than for non-readmitted patients.[3] Between 33% and 60% of discharged heart failure patients are rehospitalized within 6 months.[4, 5] Studies have found that readmission could have been prevented in at least 40% of these cases.[6] Studies of post-discharge all cause mortality rates have found that patients discharged in unstable condition are 60% more likely to die within 90 days of discharge.[7] <p>Citations for Evidence:</p> <ol style="list-style-type: none"> Krumholz, H.M., et al., Evaluating quality of care for patients with heart failure. <i>Circulation</i>,

² Citations can include, but are not limited to journal articles, reports, web pages (URLs).
NQF Measure Submission Form, V3.0

	<p>2000. 101(12): p. E122-40.</p> <p>2. Hunt, S.A., et al., 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): developed in collaboration with the American College of Chest Physicians and the International Society for Heart and Lung Transplantation: endorsed by the Heart Rhythm Society. <i>Circulation</i>, 2005. 112(12): p. e154-235.</p> <p>3. Rosenberg, A.L. and C. Watts, Patients readmitted to ICUs* : a systematic review of risk factors and outcomes. <i>Chest</i>, 2000. 118(2): p. 492-502.</p> <p>4. Gonseth, J., et al., The effectiveness of disease management programmes in reducing hospital re-admission in older patients with heart failure: a systematic review and meta-analysis of published reports. <i>Eur Heart J</i>, 2004. 25(18): p. 1570-95.</p> <p>5. Cowie, M.R., et al., Hospitalization of patients with heart failure: a population-based study. <i>Eur Heart J</i>, 2002. 23(11): p. 877-85.</p> <p>6. Hoyt, R.E. and L.S. Bowling, Reducing readmissions for congestive heart failure. <i>Am Fam Physician</i>, 2001. 63(8): p. 1593-8.</p> <p>7. Baker, D.W., et al., Trends in postdischarge mortality and readmissions: has length of stay declined too far? <i>Arch Intern Med</i>, 2004. 164(5): p. 538-44.</p>
<p>19 (1b)</p>	<p>Disparities <i>Provide evidence that demonstrates disparity in care/outcomes related to the measure focus among populations.</i></p> <p>Summary of Evidence: Studies on the relationship between race and risk of CHF readmission report mixed findings; though it appears that blacks experience lower rates of mortality for a year or more after discharge. A retrospective study conducted on a nationwide U.S. sample of 29,732 Medicare beneficiaries who were hospitalized with heart failure in 1998 and 1998 found that black patients had slightly higher adjusted risk of readmission within a year of discharge (RR, 1.09; 95% CI 1.06-1.13) but had lower crude 30 day (RR, 0.78; 95% CI 0.68-0.91) and 1 year (RR, 0.93; 95% CI, 0.88-0.98) mortality rates than white patients.[1] No significant difference in rate of readmission for CHF between blacks and whites was found in a retrospective study conducted on a cohort of 21,994 patients hospitalized with CHF in VA hospitals between 1997 and 1999. However, the authors also found that blacks were at a lower risk of both 30-day mortality after admission (OR 0.70; CI 0.60-0.92), and one year mortality after discharge (OR 0.82; CI 0.75-0.90).[2]</p> <p>Socioeconomic status may factor into readmission and mortality. A 2003 publication by the National Heart Failure project found that Medicare patients with lower SES (by zip code) had a higher adjusted risk of readmission within 1 year of discharge (RR 1.08; 95% CI 1.03-1.12) and 1 year mortality (RR 1.10; 95% CI 1.02-1.19) compared to higher SES patients.[3]</p> <p>Citations for evidence:</p> <ol style="list-style-type: none"> Rathore, S.S. and H.M. Krumholz, Race, ethnic group, and clinical research. <i>Bmj</i>, 2003. 327(7418): p. 763-4. Deswal, A., et al., Impact of race on health care utilization and outcomes in veterans with congestive heart failure. <i>J Am Coll Cardiol</i>, 2004. 43(5): p. 778-84. Rathore, S.S., et al., Socioeconomic status, treatment, and outcomes among elderly patients hospitalized with heart failure: findings from the National Heart Failure Project. <i>Am Heart J</i>, 2006. 152(2): p. 371-8.
<p>20 (1c)</p>	<p>If measuring an Outcome Describe relevance to the national health goal/priority, condition, population, and/or care being addressed: Approximately 5 million Americans have heart failure, with over 550,000 new cases diagnosed each year. Heart failure is the primary cause of death for nearly 53,000 patients annually. Heart failure is the most common primary discharge diagnosis among patients older than 65 years, accounting for 1 million hospitalizations and discharges each year. Between 33% and 60% of discharged heart failure patients are rehospitalized within 6 months. Studies have found that readmission could have been prevented in at least 40% of these cases. This indicator is intended to increase quality of care of heart failure patients during hospitalization and post discharge and decrease avoidable heart failure readmission.</p>

If not measuring an outcome, provide evidence supporting this measure topic and grade the strength of the evidence

Summarize the evidence (including citations to source) supporting the focus of the measure as follows:

- Intermediate outcome - evidence that the measured intermediate outcome (e.g., blood pressure, Hba1c) leads to improved health/avoidance of harm or cost/benefit.
- Process - evidence that the measured clinical or administrative process leads to improved health/avoidance of harm and if the measure focus is on one step in a multi-step care process, it measures the step that has the greatest effect on improving the specified desired outcome(s).
- Structure - evidence that the measured structure supports the consistent delivery of effective processes or access that lead to improved health/avoidance of harm or cost/benefit.
- Patient experience - evidence that an association exists between the measure of patient experience of health care and the outcomes, values and preferences of individuals/ the public.
- Access - evidence that an association exists between access to a health service and the outcomes of, or experience with, care.
- Efficiency- demonstration of an association between the measured resource use and level of performance with respect to one or more of the other five IOM aims of quality.

Type of Evidence Check all that apply

- | | |
|--|--|
| <input checked="" type="checkbox"/> Evidence-based guideline | <input type="checkbox"/> Quantitative research studies |
| <input type="checkbox"/> Meta-analysis | <input type="checkbox"/> Qualitative research studies |
| <input type="checkbox"/> Systematic synthesis of research | <input type="checkbox"/> Other (Please describe): |

Overall Grade for Strength of the Evidence³ (Use the USPSTF system, or if different, also describe how it relates to the USPSTF system):

Summary of Evidence (provide guideline information below):

- A meta-analysis found that the combination of comprehensive discharge planning and postdischarge support led to a 25% relative reduction in risk of readmission, 13% relative reduction in mortality, and, in some studies, improvement of quality of life scores, without increasing the cost of care.[1] Other studies, systematic reviews and meta-analyses have also found disease management programs to be effective at reducing hospital readmissions.[2-4]
- In a study examining the relationship between ACE inhibitor use and readmissions among patients with left ventricular systolic dysfunction (LVSD)-caused heart failure, patients who were not prescribed ACEIs at discharge had significantly higher rates of readmission. Among patients with no ACEI prescription, 75% had a readmission an average 138 days after initial hospitalization; 70% of patients with ACEI prescription at less-than-target doses had a readmission an average 212 days after discharge; and patients with ACEI prescription at target dose had a 60% rate of readmission 258 days after discharge.[5]
- A systematic review and meta-analysis found that educational interventions can help reduce the rate of readmission by 21%. [6] Other studies corroborate this finding.[7, 8]
- A 2004 study of the literature found little evidence to support the hypothesis that shorter length of stay is connected to either higher mortality or readmission rates.[9] In another study, longer stays were associated with higher adjusted mortality rates during the initial hospitalization and a greater adjusted risk of death post-discharge.[10]
- The Acute Decompensated Heart Failure National Registry (ADHERE) study found that only 72% of participating hospitals consistently followed ACC/AHA recommendations for patients with heart failure.[11]
- The Organized Program to Initiate Live-Saving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF) study examined the impact of current heart failure performance measures on

³The strength of the body of evidence for the specific measure focus should be systematically assessed and rated, e.g., USPSTF grading system www.ahrq.gov/clinic/uspstmeth.htm: **A** - The USPSTF recommends the service. There is high certainty that the net benefit is substantial. **B** - The USPSTF recommends the service. There is high certainty that the net benefit is moderate or there is moderate certainty that the net benefit is moderate to substantial. **C** - The USPSTF recommends against routinely providing the service. There may be considerations that support providing the service in an individual patient. There is at least moderate certainty that the net benefit is small. Offer or provide this service only if other considerations support the offering or providing the service in an individual patient. **D** - The USPSTF recommends against the service. There is moderate or high certainty that the service has no net benefit or that the harms outweigh the benefits. **I** - The USPSTF concludes that the current evidence is insufficient to assess the balance of benefits and harms of the service. Evidence is lacking, of poor quality, or conflicting, and the balance of benefits and harms cannot be determined.

patients in the first three months of discharge and found a combined mortality/readmission rate of 36.2%. The study concluded that most of the current heart failure performance measures have little relationship to patient mortality and readmission rates in the first 60-90 days post-discharge. Only prescription of an ACEI or angiotensin receptor blocker (ARB) was found to strongly influence mortality and readmission rates. The prescription of beta blockers at discharge (not currently a performance measure) was strongly associated with a reduced risk of mortality and readmission.[12]

Citations for Evidence:

1. Phillips, C.O., et al., Comprehensive discharge planning with postdischarge support for older patients with congestive heart failure: a meta-analysis. *Jama*, 2004. 291(11): p. 1358-67.
2. Gonseth, J., et al., The effectiveness of disease management programmes in reducing hospital re-admission in older patients with heart failure: a systematic review and meta-analysis of published reports. *Eur Heart J*, 2004. 25(18): p. 1570-95
3. Doughty, R.N., et al., Randomized, controlled trial of integrated heart failure management: The Auckland Heart Failure Management Study. *Eur Heart J*, 2002. 23(2): p. 139-46.
4. Roglieri, J.L., et al., Disease management interventions to improve outcomes in congestive heart failure. *Am J Manag Care*, 1997. 3(12): p. 1831-9.
5. Luthi, J.C., et al., Readmissions and the quality of care in patients hospitalized with heart failure. *Int J Qual Health Care*, 2003. 15(5): p. 413-21.
6. Gwadry-Sridhar, F.H., et al., A systematic review and meta-analysis of studies comparing readmission rates and mortality rates in patients with heart failure. *Arch Intern Med*, 2004. 164(21): p. 2315-20.
7. Doughty, R.N., et al., Randomized, controlled trial of integrated heart failure management: The Auckland Heart Failure Management Study. *Eur Heart J*, 2002. 23(2): p. 139-46.
8. Koelling, T.M., et al., Discharge education improves clinical outcomes in patients with chronic heart failure. *Circulation*, 2005. 111(2): p. 179-85.
9. Baker, D.W., et al., Trends in postdischarge mortality and readmissions: has length of stay declined too far? *Arch Intern Med*, 2004. 164(5): p. 538-44.
10. Philbin, E.F. and J.B. Roerden, Longer hospital length of stay is not related to better clinical outcomes in congestive heart failure. *Am J Manag Care*, 1997. 3(9): p. 1285-91.
11. Brackbill, M.L., R. Bashaw-Keaton, and C.S. Sytsma, Angiotensin-converting enzyme inhibitor or angiotensin receptor blocker adherence in patients with primary versus secondary diagnosis of heart failure. *Am J Manag Care*, 2007. 13(10): p. 568-70.
12. Fonarow, G.C., et al., Association between performance measures and clinical outcomes for patients hospitalized with heart failure. *Jama*, 2007. 297(1): p. 61-70.

21 (1c) **Clinical Practice Guideline** *Cite the guideline reference; quote the specific guideline recommendation related to the measure and the guideline author's assessment of the strength of the evidence; and summarize the rationale for using this guideline over others.*

Guideline Citation:

1. Hunt, S.A., et al., 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): developed in collaboration with the American College of Chest Physicians and the International Society for Heart and Lung Transplantation: endorsed by the Heart Rhythm Society. *Circulation*, 2005. 112(12): p. e154-235.
2. Bonow, R.O., et al., ACC/AHA Clinical Performance Measures for Adults with Chronic Heart Failure: a report of the American College of Cardiology/American Heart Association Task Force on Performance Measures (Writing Committee to Develop Heart Failure Clinical Performance Measures): endorsed by the Heart Failure Society of America. *Circulation*, 2005. 112(12): p. 1853-87.
3. Heart Failure in Adults. Institute for Clinical Systems Improvement, 2007.

Specific guideline recommendation:

The following are guidelines regarding treatment of heart failure. There are no specific guidelines on heart failure readmission.

	<ul style="list-style-type: none"> The American College of Cardiology/American Heart Association recommends that most heart failure patients be given a combination of diuretic, ACEI or ARB, and beta-blocker.[1] The ACC/AHA clinical performance measures for inpatients with heart failure suggest the incorporation of the following items into a patient's pre-discharge treatment: written discharge instructions (addressing activity level, diet, discharge medications, follow-up appointment, weight monitoring, and course of action if symptoms worsen); evaluation of left ventricular systolic dysfunction (LVSD); prescription of an ACEI or ARB for LVSD, unless contraindications exist; smoking cessation counseling, if applicable; and prescription of an anticoagulant for patients with chronic/recurrent atrial fibrillation.[2] The Institute for Clinical Systems Improvement (ICSI) corroborates the ACC/AHA's recommendations. Additionally, ICSI recommends the use of beta-blockers for all patients without contraindications, and aldosterone antagonists for patients with severe and debilitating (Class III and IV of the New York Heart Association classifications) heart failure.[3] <p>Guideline author's rating of strength of evidence (If different from USPSTF, also describe it and how it relates to USPSTF): A</p> <p>Rationale for using this guideline over others: ACC/AHA is a well recognize source for recommending care on heart failure.</p>
<p>22 (1c)</p>	<p>Controversy/Contradictory Evidence Summarize any areas of controversy, contradictory evidence, or contradictory guidelines and provide citations.</p> <p>Summary: None</p> <p>Citations:</p>
<p>23 (1)</p>	<p>Briefly describe how this measure (as specified) will facilitate significant gains in healthcare quality related to the specific priority goals and quality problems identified above:</p>
<p>SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES</p>	
<p>Note: Testing and results should be summarized in this form. However, additional detail and reports may be submitted as supplemental information or provided as a web page URL. If a measure has not been tested, it is only potentially eligible for time-limited endorsement.</p>	
<p>24</p>	<p>Supplemental Testing Information: attached <input type="checkbox"/> OR Web page URL:</p>
<p>25 (2b)</p>	<p>Reliability Testing</p> <p>Data/sample: We ran the claims model on 2 years of administrative claims data (2007 and 2006) from HBI commercial health plans. Year 2007 sample was used for model development and year 2006 sample was used for reliability test of the model over time.</p> <p>Analytic Method: The 30 day all cause heart failure readmission rate was the dependent variable, and the independent variable included age, gender, modified elixhauser comorbidity score, history of CHF hospitalization, chronic renal disease or dialysis, coronary artery disease, pacemaker insertion, COPD, and discharge to nursing home.</p> <p>To evaluate the model performance, we calculated several indices using the generalized linear model (GLM) with a logit link function. We calculated 3 indices for model discrimination at the patient level: the area under the receiver operating characteristic (ROC) curve, explained variation as measured by the generalized R2 statistic, and the observed outcomes in strata defined by the lowest and highest deciles based on predictive probabilities. Large values for the ROC area, R2 statistic, and a large difference in predicted probabilities between highest and lowest deciles provide evidence that the model has good discrimination. We also calculated 2 indices for model calibration: Hosmer and Lemeshow Goodness-of-Fit test, and the global null hypothesis test of betas (likelihood ratio test). Non-significant p-value by Hosmer and Lemeshow and significant p-value by likelihood ratio test provide evidence that the model has good calibration.</p> <p>Finally, we used the hierachial generalized linear model (i.e.random intercept model) to re-estimate the</p>

	<p>coefficients to account for the natural clustering of observations within hospitals.</p> <p>Testing Results:</p> <table border="1"> <thead> <tr> <th>Model</th> <th>Sample size</th> <th>Crude rate</th> <th>Adjusted R²</th> <th>Predictability low-high decile</th> <th>ROC curve Area</th> <th>Hosmer & Lemeshow Goodness of fit test</th> <th>Wald Chi-square p value</th> </tr> </thead> <tbody> <tr> <td>Derivation sample</td> <td>6707</td> <td>0.211</td> <td>0.044</td> <td>12.5% - 37.2%</td> <td>0.618</td> <td>0.7041</td> <td><0.0001</td> </tr> <tr> <td>Reliability test sample</td> <td>3328</td> <td>0.154</td> <td>0.052</td> <td>5.7% - 31.8%</td> <td>0.635</td> <td>0.6453</td> <td><0.0001</td> </tr> </tbody> </table>	Model	Sample size	Crude rate	Adjusted R ²	Predictability low-high decile	ROC curve Area	Hosmer & Lemeshow Goodness of fit test	Wald Chi-square p value	Derivation sample	6707	0.211	0.044	12.5% - 37.2%	0.618	0.7041	<0.0001	Reliability test sample	3328	0.154	0.052	5.7% - 31.8%	0.635	0.6453	<0.0001
Model	Sample size	Crude rate	Adjusted R ²	Predictability low-high decile	ROC curve Area	Hosmer & Lemeshow Goodness of fit test	Wald Chi-square p value																		
Derivation sample	6707	0.211	0.044	12.5% - 37.2%	0.618	0.7041	<0.0001																		
Reliability test sample	3328	0.154	0.052	5.7% - 31.8%	0.635	0.6453	<0.0001																		
26	<p>Validity Testing</p> <p>(2c) Data/sample: Phametrics Medicare data with patients older than 65 was used as validation sample</p> <p>Analytic Method: See analytic method described in box 25 (above).</p> <p>Testing Results:</p> <table border="1"> <thead> <tr> <th>Model</th> <th>Sample size</th> <th>Crude rate</th> <th>Adjusted R²</th> <th>Predictability low-high decile</th> <th>ROC curve Area</th> <th>Hosmer & Lemeshow Goodness of fit test</th> <th>Wald Chi-square p value</th> </tr> </thead> <tbody> <tr> <td>Derivation sample</td> <td>6707</td> <td>0.211</td> <td>0.044</td> <td>12.5% - 37.2%</td> <td>0.618</td> <td>0.7041</td> <td><0.0001</td> </tr> <tr> <td>Validation sample</td> <td>19533</td> <td>0.135</td> <td>0.188</td> <td>5.0% - 42.4%</td> <td>0.741</td> <td>0.5303</td> <td><0.0001</td> </tr> </tbody> </table> <p>Validity of heart failure readmission crude rate: - See results presented in the above table and in box 25 (above). The crude rate obtained from the commercial plans across 2 years are internally consistent, and consistent with data presented by the NQF endorsed CMS 30-day heart failure readmission measure. During on testing, we found 30-day readmission rate ranged from 13.5% to 21.1%. CMS found the following results for 30-day readmission rate:</p> <ol style="list-style-type: none"> (1) 1998 derivation sample (ResDAC dat) - 15.5%, (2) 1998 validation sample (ResDAC data) - 15.6% (3) 1999 validation sample (RTI data) - 18.8% (4) 2000 validation sample (RTI data) - 18.7% (5) 2001 validation sample (RTI data) - 17.8%. <p>Validity of risk-adjustment model: - See result presented in box 30 (below). The model performed consistently across the derivation sample and the validation samples. Statistically significant covariates predicted readmission in a manner consistent with our conceptual model. For example we hypothesized that members with greater comorbidity or discharged to the nursing home would be more likely to be readmitted based on conceptual model, and our results are consistent with this hypothesis.</p>	Model	Sample size	Crude rate	Adjusted R ²	Predictability low-high decile	ROC curve Area	Hosmer & Lemeshow Goodness of fit test	Wald Chi-square p value	Derivation sample	6707	0.211	0.044	12.5% - 37.2%	0.618	0.7041	<0.0001	Validation sample	19533	0.135	0.188	5.0% - 42.4%	0.741	0.5303	<0.0001
Model	Sample size	Crude rate	Adjusted R ²	Predictability low-high decile	ROC curve Area	Hosmer & Lemeshow Goodness of fit test	Wald Chi-square p value																		
Derivation sample	6707	0.211	0.044	12.5% - 37.2%	0.618	0.7041	<0.0001																		
Validation sample	19533	0.135	0.188	5.0% - 42.4%	0.741	0.5303	<0.0001																		
27	<p>Measure Exclusions <i>Provide evidence to justify exclusion(s) and analysis of impact on measure results during testing.</i></p> <p>(2d) Summary of Evidence supporting exclusion(s): We excluded patients in hospice during the 356 days prior to index date (date of discharge) through 30 days after index date, and members who expired during the index hospitalization. It is evident that members who expired during the index hospitalization should be excluded because they have no chance of being re-admitted for heart failure. We conceptually excluded patients who receive hospice on discharge because members on hospice most likely have end stage heart failure and would be admitted only for palliative treatment. Less than 1% of our commercially insured sample received hospice during the time period specified and including and excluding these patients made</p>																								

	<p>no difference in the results.</p> <p>Citations for Evidence:</p> <p>Data/sample:</p> <p>Analytic Method:</p> <p>Testing Results:</p>																																																																																					
28 (2e)	<p>Risk Adjustment Testing Summarize the testing used to determine the need (or no need) for risk adjustment and the statistical performance of the risk adjustment method.</p> <p>Data/sample: Data from one comercial health plan in the derivation sample</p> <p>Analytic Method: We used the HBI heart-failure readmission model to calculate the predicted rate of hospitals in the selected health plan. A threshold of 10 was applied to exclude hospitals with fewer than 10 heart failure patients. We then ranked the hospitals by the predicted rate: These results are presented in the table below. Note that lower predicted rates imply better performance.</p> <p>Testing Results:</p> <table border="1"> <thead> <tr> <th>Hospital Ranking</th> <th># of Cases</th> <th># of Readmission</th> <th>Predicted Rate</th> <th>Crude Rate</th> </tr> </thead> <tbody> <tr><td>HOSP 01</td><td>129</td><td>15</td><td>0.123</td><td>0.116</td></tr> <tr><td>HOSP 02</td><td>69</td><td>6</td><td>0.124</td><td>0.087</td></tr> <tr><td>HOSP 03</td><td>158</td><td>19</td><td>0.125</td><td>0.120</td></tr> <tr><td>HOSP 04</td><td>124</td><td>15</td><td>0.130</td><td>0.121</td></tr> <tr><td>HOSP 05</td><td>22</td><td>2</td><td>0.137</td><td>0.091</td></tr> <tr><td>HOSP 06</td><td>33</td><td>3</td><td>0.139</td><td>0.091</td></tr> <tr><td>HOSP 07</td><td>74</td><td>11</td><td>0.141</td><td>0.149</td></tr> <tr><td>HOSP 08</td><td>50</td><td>7</td><td>0.144</td><td>0.140</td></tr> <tr><td>HOSP 09</td><td>14</td><td>3</td><td>0.150</td><td>0.214</td></tr> <tr><td>HOSP 10</td><td>28</td><td>5</td><td>0.151</td><td>0.179</td></tr> <tr><td>HOSP 11</td><td>32</td><td>6</td><td>0.151</td><td>0.188</td></tr> <tr><td>HOSP 12</td><td>51</td><td>11</td><td>0.154</td><td>0.216</td></tr> <tr><td>HOSP 13</td><td>119</td><td>21</td><td>0.155</td><td>0.176</td></tr> <tr><td>HOSP 14</td><td>99</td><td>17</td><td>0.158</td><td>0.172</td></tr> <tr><td>HOSP 15</td><td>172</td><td>33</td><td>0.159</td><td>0.192</td></tr> <tr><td>HOSP 16</td><td>107</td><td>28</td><td>0.195</td><td>0.262</td></tr> </tbody> </table> <p>The risk-adjustment made a difference in comparing hospital performance. Hospital 01 has the lowest predicted heart-failure readmission rate, but higher crude rates than hospital 2, 5, and 6. In other words, the rankings of the hospitals are changed by applying the risk-adjustment model.</p> <p>► If outcome or resource use measure not risk adjusted, provide rationale:</p>	Hospital Ranking	# of Cases	# of Readmission	Predicted Rate	Crude Rate	HOSP 01	129	15	0.123	0.116	HOSP 02	69	6	0.124	0.087	HOSP 03	158	19	0.125	0.120	HOSP 04	124	15	0.130	0.121	HOSP 05	22	2	0.137	0.091	HOSP 06	33	3	0.139	0.091	HOSP 07	74	11	0.141	0.149	HOSP 08	50	7	0.144	0.140	HOSP 09	14	3	0.150	0.214	HOSP 10	28	5	0.151	0.179	HOSP 11	32	6	0.151	0.188	HOSP 12	51	11	0.154	0.216	HOSP 13	119	21	0.155	0.176	HOSP 14	99	17	0.158	0.172	HOSP 15	172	33	0.159	0.192	HOSP 16	107	28	0.195	0.262
Hospital Ranking	# of Cases	# of Readmission	Predicted Rate	Crude Rate																																																																																		
HOSP 01	129	15	0.123	0.116																																																																																		
HOSP 02	69	6	0.124	0.087																																																																																		
HOSP 03	158	19	0.125	0.120																																																																																		
HOSP 04	124	15	0.130	0.121																																																																																		
HOSP 05	22	2	0.137	0.091																																																																																		
HOSP 06	33	3	0.139	0.091																																																																																		
HOSP 07	74	11	0.141	0.149																																																																																		
HOSP 08	50	7	0.144	0.140																																																																																		
HOSP 09	14	3	0.150	0.214																																																																																		
HOSP 10	28	5	0.151	0.179																																																																																		
HOSP 11	32	6	0.151	0.188																																																																																		
HOSP 12	51	11	0.154	0.216																																																																																		
HOSP 13	119	21	0.155	0.176																																																																																		
HOSP 14	99	17	0.158	0.172																																																																																		
HOSP 15	172	33	0.159	0.192																																																																																		
HOSP 16	107	28	0.195	0.262																																																																																		
29 (2g)	<p>Testing comparability of results when more than 1 data method is specified (e.g., administrative claims or chart abstraction)</p> <p>Data/sample: N/A - only administrative model was developed.</p> <p>Analytic Method:</p> <p>Results:</p>																																																																																					
30 (2f)	<p>Provide Measure Results from Testing or Current Use Results from testing</p> <p>Data/sample: We derived the claims model using year 2007 administrative data from HBI comercial health plans, we conducted reliablity test over time using year 2006 administrative data from HBI comercial health plans, and we further validated the model using Pharmetrics Medicare data with</p>																																																																																					

patients older than 65 years.

Methods to identify statistically significant and practically/meaningfully differences in performance:

The 30 day all cause heart failure readmission rate was the dependent variable, and the independent variables included age, gender, elixhauser comorbidity score, history of CHF hospitalization in the past year, chronic renal disease or dialysis, coronary artery disease, pacemaker insertion, COPD, and discharge to nursing home. We used the hierachial generalized linear model (i.e. random intercept model) to estimate the coefficients to account for the natural clustering of observations within hospitals. The coefficient estimate, odds ratio, p-value, and 95% confidence interval of the estimate were provided in the following for HBI 2007 derivation model, HBI year 2006 reliability test model, and Pharmetrics Medicare validation model.

Results:

Derivation model:

Year 2007 Administrative Claims Hierarchical Model (n=6707)

Variable	estimate	Odds ratio	p-value	95% CI
Intercept	-3.38	0.03	0.00	0.01-0.13
Age	0.33	1.40	0.09	0.95-2.06
Age squared	-0.02	0.98	0.16	0.95-1.01
Female	0.00	1.00	0.95	0.88-1.13
Elixhauser comorbidity scores	0.10	1.10	0.00	1.07-1.14
History of CHF	0.28	1.32	0.00	1.11-1.59
Chronic Renal disease or dialysis	0.24	1.27	0.00	1.12-1.45
Coronary Artery Disease	0.19	1.21	0.01	1.05-1.39
Pacemaker insertion	0.08	1.08	0.51	0.86-1.36
COPD	0.06	1.07	0.63	0.82-1.38
Discharge to nursing home	0.65	1.92	0.00	1.64-2.24

Reliability test model:

Year 2006 Administrative Claims Hierarchical Model (n=3328)

Variable	estimate	Odds ratio	p-value	95% CI
Intercept	-3.08	0.05	0.00	0.01-0.25
Age	0.14	1.15	0.60	0.69-1.93
Age squared	-0.01	0.99	0.54	0.95-1.03
Female	0.05	1.05	0.62	0.87-1.27
Elixhauser comorbidity scores	0.14	1.15	0.00	1.10-1.20
History of CHF	0.13	1.14	0.36	0.86-1.52
Chronic Renal disease or dialysis	0.17	1.18	0.12	0.96-1.46
Coronary Artery Disease	0.35	1.41	0.00	1.13-1.77
Pacemaker insertion	0.00	1.00	0.98	0.71-1.40
COPD	0.60	1.82	0.00	1.24-1.68
Discharge to nursing home	0.68	1.97	0.00	1.40-2.76

Validation Model:

Pharmetrics Administrative Claims Hierarchical Model (n=19553)

Variable	estimate	Odds ratio	p-value	95% CI
Intercept	2.07	7.90	0.20	0.34-182
Age	-1.22	0.30	0.00	0.14-0.63
Age squared	0.07	1.08	0.00	1.03-1.13
Female	0.13	1.14	0.01	1.03-1.25
Elixhauser comorbidity scores	0.02	1.02	0.11	1.00-1.04
History of CHF	0.12	1.13	0.93	0.09-13.7
Chronic Renal disease or dialysis	0.47	1.60	0.00	1.44-1.77

	<p>Coronary Artery Disease 0.18 1.20 0.00 1.09-1.33</p> <p>Pacemaker insertion 0.11 1.11 0.25 0.92-1.34</p> <p>COPD 0.28 1.33 0.00 1.14-1.54</p> <p>Discharge to nursing home 2.02 7.55 0.00 6.87-8.30</p> <p>Of note all the age variables are measured in 10 years of age.</p> <p>Statistically significant covariates predicted readmission in a manner consistent with our conceptual model. In general, members with greater comorbidity, history of congestive heart failure, chronic renal disease, coronary artery disease, pacemaker insertion, and COPD, and who were discharged to nursing home after acute hospital care were more likely to be readmitted to the hospital in 30 days.</p>
31 (2h)	<p>Identification of Disparities</p> <p>► If measure is stratified by factors related to disparities (i.e. race/ethnicity, primary language, gender, SES, health literacy), provide stratified results: Not stratified</p> <p>► If disparities have been reported/identified, but measure is not specified to detect disparities, provide rationale:</p>
USABILITY	
32 (3)	<p><i>Current Use Testing completed</i> If in use, how widely used (select one) ► If "other," please describe:</p> <p><input type="checkbox"/> Used in a public reporting initiative, name of initiative: Sample report attached <input type="checkbox"/> OR Web page URL:</p>
33 (3a)	<p>Testing of Interpretability (<i>Testing that demonstrates the results are understood by the potential users for public reporting and quality improvement</i>)</p> <p>Data/sample:</p> <p>Methods:</p> <p>Results:</p>
34 (3b, 3c)	<p>Relation to other NQF-endorsed™ measures</p> <p>► Is this measure similar or related to measure(s) already endorsed by NQF (on the same topic or the same target population)? <i>Measures can be found at www.qualityforum.org under Core Documents.</i></p> <p>Check all that apply</p> <p><input type="checkbox"/> Have not looked at other NQF measures <input type="checkbox"/> Other measure(s) on same topic</p> <p><input checked="" type="checkbox"/> Other measure(s) for same target population <input type="checkbox"/> No similar or related measures</p> <p>Name of similar or related NQF-endorsed™ measure(s): HF 30-Day Readmission Administrative Data Model (authored by CMS)</p> <p>Are the measure specifications harmonized with existing NQF-endorsed™ measures? Yes, fully harmonized</p> <p>► If not fully harmonized, provide rationale:</p> <p>Describe the distinctive, improved, or additive value this measure provides to existing NQF-endorsed measures:</p> <p>The CMS 30-day readmission model measures 30-day all-cause adjusted readmission rate for members 65 years and older, and was tested on Medicare data only. HBI heart-failure readmission model measures 30-day all-cause adjusted readmission rate for members 19 years and older and was tested on commercial data including Medicare commercial data.</p> <p>This model is similar to the CMS model in some respects. We tested many of the same covariates</p>

	<p>concepts tested in the CMS model, and added a few covariates not in the CMS model which have been found to be important in the literature.</p> <p>Similarities to CMS model:</p> <p>In the development of the CMS model CMS tested the following candidate covariates concepts which we also tested: cardiovascular disease, renal failure, dialysis, COPD, diabetes, dementia, and Charlson Comorbidity Index (Deyo score, modified). We combined renal failure and dialysis variable for the commercially insured for parsimony. Final CMS model included renal failure, COPD, and coronary artery disease. This final CMS model was based on step-wise regression results with exit criterion of 0.0001 ran solely on CMS data set. Dialysis and Charlson Comorbidity Index did not meet the final cut-off criteria based on testing using CMS data.</p> <p>The first difference between the endorsed CMS model and the HBI model is that the HBI model has been developed and tested for members younger than 65 years of age with heart failure.</p> <p>Commerical data sets revealed that at approximately a third of heart failure patients were under 65. The HBI specific commerical data set (8 health plan, approximately 20 million lives) revealed 56% of the patients with at least one hospital admission for heart failure were under 65 years of age in the 2006, and 28% were under 65 years of age in 2007. The percentage of patients under 65 years decreased because in 2007, some health plan were able to obtain secondary payer data information (e.g., Medicare Advantage data). The Pharmedics specific data set with 98 health plans and 62 million lives revealed that 31% of the patients with at least one hospital admission for heart failure were under 65 years of age.</p> <p>Thus, in the CMS model (because of the limitation of their data set) age is used as a linear variable. We found that when we ran age as a linear variable, our model had worse performance than if we had run age as non-linear variables (e.g., age and age²). Our hypothesis is that commercial plans have ~ 30% of the patients younger than 65 years of age who were admitted for CHF. The linear age variable does not fit commercial data well because it has a broader age range.</p> <p>Second, we tested covariates not present in CMS model conceptually but were important in the literature (e.g., discharge to nursing home, history of CHF admission). In particular, administrative claims data do not capture direct information regarding severity of heart failure. Thus, we used discharged to nursing home as proxy measure for more severe heart failure. In our testing, heart failure patients who were discharge to nursing home were significantly more likely to be re-admitted to the hospital in 30 days. This variable have very consistent and signifcant performance in both the derivation and validation data sets. This variable is a critical contributor to our model performance.</p> <p>CMS admitted that their 30-day readmission model had poor performance, with an R2 ranging from 0.01 to 0.02, ROC area ranging from 0.55 to 0.59, and low predictive ability. In contrast, the HBI heart-failure readmission model has good model performance, with an R2 ranging from 0.05 to 0.18, and ROC area ranging from 0.62 to 0.74. In addition, the CMS heart failure re-admission model requires 24 covariates and is most appropriately run on sample size > 240 heart failure patients. As a result, the CMS model cannot be used by smaller hospitals and be used as a quality indicator to score physician or physician group performance. In contrast, the HBI model is a much more parsimonious model and requires only 8 covariates and can be run on sample sizes of > 80 heart failure patients, and is therefore more accesible for scoring smaller hospitals and even large physician groups.</p> <p>Although model performance metrics will vary based on the dataset used and can't be easily compared unless the models were applied to the same dataset, our hypothesis that HBI model may performed better than CMS model because we used age in a quadratic and non-linear fashion and we included a proxy for severity of illness (i.e., discharge to nursing home).</p>
FEASIBILITY	
<p>35 (4a)</p>	<p><i>How are the required data elements generated? Check all that apply</i></p> <p><input checked="" type="checkbox"/> <i>Data elements are generated concurrent with and as a byproduct of care processes during care delivery (e.g., blood pressure or other assessment recorded by personnel conducting the assessment)</i></p>

	<input type="checkbox"/> Data elements are generated from a patient survey (e.g., CAHPS) <input type="checkbox"/> Data elements are generated through coding performed by someone other than the person who obtained the original information (e.g., DRG or ICD-9 coding on claims) <input type="checkbox"/> Other, Please describe:
36 (4b)	<p>Electronic Sources All data elements</p> <p>► If all data elements are not in electronic sources, specify the near-term path to electronic collection by most providers:</p> <p>► Specify the data elements for the electronic health record: ICD-9 diagnosis codes, ICD-9 Proc Codes, CPT-4 codes, HCPCS codes, UB revenue codes, NDC code, DRG codes</p>
37 (4c)	<p>Do the specified exclusions require additional data sources beyond what is required for the other specifications? <i>No</i></p> <p>► If yes, provide justification:</p>
38 (4d)	<p>Identify susceptibility to inaccuracies, errors, or unintended consequences of the measure: <i>It is possible that patients coded as having heart failure at discharge do not have heart failure and have been inaccurately coded. However, previous studies have shown that any given heart failure claim has a 93% specificity of identifying patients with heart failure, and a heart failure face-to-face claim (paired with evaluation and management code as it is defined in HBI denominator algorithm) has a 96% specificity in identifying patients with heart failure. (Rector et al., 2004)</i></p> <p><i>Describe how could these potential problems be audited: HBI has developed an online tool (currently in use by several health plans), which allows hospitals or physicians the opportunity to supplement information through self-report via a secured web site. Via this website, physicians or hospitals are able to identify specific patients during the measurement period and who reportedly were admitted for heart failure to verify the diagnosis. The physician or hospital administrator can then manually enter corrections to the patient record via the website with the understanding that the information entered is subject to clinical review. The hybrid quality score (via administrative claims and self report) can be updated on a quarterly basis.</i></p> <p><i>Did you audit for these potential problems during testing? No If yes, provide results:</i></p>
39 (4e)	<p>Testing feasibility Describe what have you learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data/missing data, timing/frequency of data collection, patient confidentiality, time/cost of data collection, other feasibility/ implementation issues:</p>
CONTACT INFORMATION	
40	<p>Web Page URL for Measure Information Describe where users (implementers) should go for more details on specifications of measures, or assistance in implementing the measure. Web page URL: <i>N/A</i></p>
41	<p>Measure Intellectual Property Agreement Owner Point of Contact First Name: <i>Zak</i> MI: Last Name: <i>Ramadan-Jradi</i> Credentials (MD, MPH, etc.): <i>MD, MPH</i> Organization: <i>Health Benchmarks®</i> Street Address: <i>21650 Oxnard St., Suite 550</i> City: <i>Woodland Hills</i> State: <i>CA</i> ZIP: <i>91367-7806</i> Email: <i>zramadan@us.imshealth.com</i> Telephone: <i>818-676-2820</i> ext:</p>
42	<p>Measure Submission Point of Contact If different than IP Owner Contact First Name: <i>Karen</i> MI: Last Name: <i>Hsu</i> Credentials (MD, MPH, etc.): <i>MPH, MBA</i> Organization: <i>Health Benchmarks®</i> Street Address: <i>21650 Oxnard St., Suite 550</i> City: <i>Woodland Hills</i> State: <i>CA</i> ZIP: <i>91367-7806</i> Email: <i>khsu@us.imshealth.com</i> Telephone: <i>541-550-7983</i> ext:</p>
43	<p>Measure Developer Point of Contact If different than IP Owner Contact First Name: <i>Judy</i> MI: <i>Y</i> Last Name: <i>Chen</i> Credentials (MD, MPH, etc.): <i>MD, MSHS</i> Organization: <i>Health Benchmarks®</i></p>

	Street Address: 21650 Oxnard St., Suite 550 City: Woodland Hills State: CA ZIP: 91367-7806 Email: judy.chen@us.imshealth.com Telephone: 818-676-2883 ext:
44	<p>Measure Steward Point of Contact If different than IP Owner Contact</p> <p><i>Identifies the organization that will take responsibility for updating the measure and assuring it is consistent with the scientific evidence and current coding schema; the steward of the measure may be different than the developer.</i></p> <p>First Name: MI: Last Name: Credentials (MD, MPH, etc.): Organization: Street Address: City: State: ZIP: Email: Telephone: ext</p>
ADDITIONAL INFORMATION	
45	<p>Workgroup/Expert Panel involved in measure development No workgroup or panel used</p> <p>▶ If workgroup used, describe the members' role in measure development: ▶ Provide a list of workgroup/panel members' names and organizations:</p>
46	<p><i>Measure Developer/Steward Updates and Ongoing Maintenance</i></p> <p><i>Year the measure was first released: 2008</i></p> <p><i>Month and Year of most recent revision: October, 2008</i></p> <p><i>What is the frequency for review/update of this measure? Annually</i></p> <p><i>When is the next scheduled review/update for this measure? September, 2009</i></p>
47	<p>Copyright statement/disclaimers:</p> <p>© 2008 Health Benchmarks® Confidential and Proprietary All Rights Reserved</p>
48	Additional Information:
49	I have checked that the submission is complete and any blank fields indicate that no information is provided. <input checked="" type="checkbox"/>
50	Date of Submission (MM/DD/YY): 10/31/08