AHRQ Quality Indicators

Inpatient Quality Indicators Composite Measure Workgroup

Final Report

Agency for Healthcare Research and Quality
Advancing Excellence in Health Care • www.ahrq.gov

Agency for Healthcare Research and Quality
Quality Indicators (AHRQ QI)

Inpatient Quality Indicators (IQI)
Composite Measure Workgroup
Final Report
March 2008

Table of Contents

AHRQ Quality Indicators
Inpatient Quality Indicators
Composite Measure Workgroup
Final Report

1. Introduction

Many users of the AHRQ Quality Indicators (AHRQ QIs) have expressed interest in the development of one or more composite measures. In particular, the National Healthcare Quality Report and the National Healthcare Disparities Report[1] staff asked the AHRQ QI program to develop composite measures for use in these reports. A composite measure for the Prevention Quality Indicators was developed initially.[2] The goal of the development effort was to develop a composite measure that might be used to monitor performance over time or across regions and populations using a methodology that applied at the national, regional, State, or provider/area level. This report describes the construction of two composite measures for the Inpatient Quality Indicators (IQIs): *Mortality for Selected Procedures* and *Mortality for Selected Conditions*.

To assist in the development of a composite measure methodology, the AHRQ QI Composite Measure Workgroup held several conference calls to discuss important issues and considerations and to provide feedback on preliminary results. To maintain the focus on the general composite measure methodology, the Workgroup did not consider the merits of including individual indicators in the composites. Rather, all available Inpatient Quality Indicators that met the conceptual criteria were included. The members of the AHRQ QI Composite Measure Workgroup are listed in Appendix A.

This report is very technical in nature. To facilitate future use of the composite, the AHRQ QI program plans to develop more accessible explanatory narrative on the composite measures as part of the reporting template initiative.

For more information on the Inpatient Quality Indicators, including the selection criteria, coding, and specifications, see the Guide to Inpatient Quality Indicators and the Inpatient Quality Indicators Technical Specifications, available on the AHRQ QI Web site (http://qualityindicators.ahrq.gov).[3]

2. Reasons for Composite Measures

Before considering alternative approaches to composite measures, one might consider why composite measures are potentially useful and for what purpose.

2.1. Benefits of Composite Measures

Composite measures have several potential benefits over individual indicators:

- *Summarize quality across multiple indicators*. There are 30 provider-level IQIs for various conditions and procedures, making it difficult to formulate general statements about overall trends or differences in quality.

---

[1] The most recent National Healthcare Quality Report and National Healthcare Disparities Report may be found at http://www.ahrq.gov/qual/measurix.htm.
[2] A report describing the composite measure for the Prevention Quality Indicators can be found at: http://www.qualityindicators.ahrq.gov/downloads/technical/AHRQ QI PQI Composite Report Final.pdf.
[3] Guide: http://www.qualityindicators.ahrq.gov/downloads/iqi/iqi_guide_v31.pdf; Technical Specifications: http://www.qualityindicators.ahrq.gov/downloads/iqi/iqi_technical_specs_v31.pdf.

- *Improve ability to detect quality differences*. Combining information from multiple indicators may result in greater discrimination in performance than is evident from individual indicators.

- *Identify important domains and drivers of quality*. To the extent that certain indicators track together, or track with certain process or structural characteristics of providers, one may identify the important domains and drivers of quality.

- *Prioritize action for quality improvement*. Individual indicators that contribute a larger share to the composite may be targets for quality improvement activity.

- *Make current decisions about future (unknown) health care needs*. Depending on how the component indicators are weighted, composites may reflect the likely health outcomes for an individual or population.

- *Avoid cognitive "shortcuts."* Research suggests that individuals faced with too many factors in making a decision take cognitive shortcuts that might not be in their best interest. Composites may help to ensure that decisions are made appropriately.

2.2. Concerns About Composite Measures

Despite these benefits, there are concerns about using composite measures, depending on how the composite measure is constructed:

- *Can mask important differences and relationships among components*. Composite measures might mask the fact that two components are inversely related, or an "average" provider might be high on one component and low on another.

- *May not be actionable*. It might not be clear what action a provider should take given high or low performance on a composite measure.

- *May not identify which parts of the health care system contribute most to quality*. To the extent that the composite is not connected to the interventions important for the component measures, it might be difficult to know how the composite contributes to improved quality.

- *Can detract from the impact and credibility of reports*. The composite measure might not reflect the evidence base of the component indicators.

2.3. Potential Uses of Composite Measures

Composite measures have many potential uses:

- *Consumers* might use composite measures to select a hospital or health plan either before or after a health event.

- *Providers* might use composite measures to identify the domains and drivers of quality.

- *Purchasers* might use composite measures to select hospitals or health plans in order to improve the health of employees.

- *Policymakers* might use composite measures to set policy priorities in order to improve the health of a population.

3. Alternative Perspectives on Composite Measures

Two alternative perspectives on composite measures guide the development of a composite measure methodology:

- *Signaling perspective*, which seeks to guide decisionmaking by providing information that will result in actions leading to some intended result. The ultimate evaluation criterion for the composite measure is the usefulness of the measure for achieving the intended result. An example of a composite measure reflecting the signaling perspective is the Dow Jones Industrial Average used to guide decisionmaking on allocating investment resources.

- *Psychometric perspective*, which seeks to capture an underlying construct of quality based on multiple single indicators. The ultimate evaluation criterion for the composite measure is the extent to which the components reflect that construct. An example of a composite measure reflecting the psychometric perspective is the IQ test used to capture a construct labeled "intelligence."

The methodology used for the AHRQ QI composite measures reflects the signaling perspective, in that the primary intent of the measures is to guide decisionmaking in terms of where to allocate resources to improve quality rather than to capture an underlying construct of quality.

4. Methodology for the AHRQ QI Composite Measures

4.1. Composite Measure Development Criteria

This report describes the construction of two composite measures for the IQIs: *Mortality for Selected Procedures* and *Mortality for Selected Conditions*. Appendix B presents IQI composite tables (Tables 1-9). Table 1 shows the reference population, including the mortality rate for each procedure or condition.

The basic criteria used to guide the development of the methodology were:

- *Evidence based*. The composite measure should be based on indicator components that are important, reliable, valid, and minimally biased.

- *Conceptually coherent*. The components of the composite measure should be related to one another conceptually.

- *Empirically coherent*. The components of the composite measure should be related to one another empirically.

- *Intended use*. The composite measures should be constructed in a manner appropriate to the intended use, whether that is comparative reporting or quality improvement.

Applying these criteria to the IQIs, one conceptually coherent grouping of the indicators was to include the mortality for selected procedure indicators in one composite and the mortality for selected condition indicators in another composite (see table below), primarily because the former indicators apply to a smaller group of hospitals than the latter (see Table 2) and have a similar quality rationale. In addition, this grouping was generally empirically coherent as the measures tend to be positively correlated with one another, although not strongly so, especially in the case of the procedure indicators (see Tables 3 and 4).

AHRQ IQI Composite Measures

Mortality for Selected Procedures	Mortality for Selected Conditions
IQI #08 In-Hosp Mort Esophageal Resection IQI #09 In-Hosp Mort Pancreatic Resection IQI #11 In-Hosp Mort AAA Repair IQI #12 In-Hosp Mort CABG IQI #13 In-Hosp Mort Craniotomy IQI #14 In-Hosp Mort Hip Replacement IQI #30 In-Hosp Mort PTCA IQI #31 In-Hosp Mort Carotid Endarterectomy	IQI #15 In-Hosp Mort AMI IQI #16 In-Hosp Mort CHF IQI #17 In-Hosp Mort Stroke IQI #18 In-Hosp Mort GI Hemorrhage IQI #19 In-Hosp Mort Hip Fracture IQI #20 In-Hosp Mort Pneumonia

Note: AAA=abdominal aortic aneurysm; CABG=coronary artery bypass graft; PTCA=percutaneous transluminal coronary angioplasty; AMI=acute myocardial infarction; CHF=congestive heart failure; GI=gastrointestinal.

4.2. AHRQ QI Composite Measure Methodology

The general methodology for the AHRQ QI composite measures might be described as constructing a "composite of composites." The first "composite" is the reliability-adjusted ratio, which is a weighted average of the risk-adjusted ratio and the reference population ratio, where the weight is determined empirically. The second "composite" is a weighted average of the component indicators, where the weights are selected based on the intended use of the composite measure. These weights might be determined empirically or based on nonempirical considerations.

4.3. Construction of AHRQ QI Composite Measure

The basic steps for computing the composite follow.

Step 1. Compute the risk-adjusted rate and confidence interval

The AHRQ QI risk-adjusted rate is computed based on a simple logistic regression model[4] for calculating a predicted value for each case. Then the predicted values among all the cases in the hospital are summed to compute the expected rate. The risk-adjusted rate is computed using indirect standardization as the observed rate (OR) divided by the expected rate (ER), with the result multiplied by the reference population rate: (RR) = (OR/ER × PR). The reference population used in this analysis includes the States participating in the Healthcare Cost &

---

[4] Release 3.1 (fiscal year 2007) of the AHRQ QI software adopted a hierarchical modeling methodology for the risk adjustment, but the composite methodology remains the same.

Utilization Project (HCUP) for 2001-2003, consisting of 38 States and approximately 90 million discharges.[5]

Step 2. Scale the risk-adjusted rate using the reference population

Table 1 shows the reference population numerator, denominator, and rate for each IQI. The relative magnitudes of the rates vary from indicator to indicator. To combine the component indicators using a common scale, each indicator's risk-adjusted rate is divided by the reference population rate to yield a ratio. The components of the composite are therefore defined in terms of a ratio to the reference population rate for each indicator. The component indicators are scaled by the reference population rate so that each indicator reflects the degree of deviation from the overall average performance.

Step 3. Compute the reliability-adjusted ratio

The reliability-adjusted ratio (RAR) is computed as the weighted average of the risk-adjusted ratio and the reference population ratio, where the weights vary from 0 to 1, depending on the degree of reliability for the indicator and provider (or other unit of analysis).

$$\text{RAR} = [\text{risk-adjusted ratio} \times \text{weight}] + [\text{reference population ratio} \times (1 - \text{weight})]$$

Table 5 shows the average reliability weights for the IQIs based on denominator size. For small providers, the weight is closer to 0. For large providers, the weight is closer to one. For a given provider, if the denominator is zero, then the weight assigned is zero (i.e., the reliability-adjusted ratio is the reference population ratio).

Step 4. Select the component weights

The composite measure is the weighted average of the scaled and reliability-adjusted ratios for the component indicators. Table 6 shows examples of alternative weights that might be used. Other weights are also possible.

Single indicator weight. In this case, the composite is simply the reliability-adjusted ratio for a single indicator. For the procedure indicators, the single indicator weight composite is a "volume-outcome" composite since the reference population rate varies with the volume of the

---

[5] The State data organizations that participated in the 2001-2003 HCUP State Inpatient Databases are: Arizona Department of Health Services; California Office of Statewide Health Planning and Development; Colorado Health and Hospital Association; Connecticut - Chime, Inc.; Florida Agency for Health Care Administration; Georgia - AHA: An Association of Hospitals and Health Systems; Hawaii Health Information Corporation; Illinois Health Care Cost Containment Council; Indiana Hospital & Health Association; Iowa Hospital Association; Kansas Hospital Association; Kentucky Department for Public Health; Maine Health Data Organization; Maryland Health Services Cost Review Commission; Massachusetts Division of Health Care Finance and Policy; Michigan Health & Hospital Association; Minnesota Hospital Association; Missouri Hospital Industry Data Institute; Nebraska Hospital Association; Nevada Department of Human Resources; New Hampshire Department of Health & Human Services; New Jersey Department of Health and Senior Services; New York State Department of Health; North Carolina Department of Health and Human Services; Ohio Hospital Association; Oregon Association of Hospitals and Health Systems; Pennsylvania Health Care Cost Containment Council; Rhode Island Department of Health; South Carolina Budget & Control Board; South Dakota Association of Healthcare Organizations; Tennessee Hospital Association; Texas Health Care Information Council; Utah Department of Health; Vermont Association of Hospitals and Health Systems; Virginia Health Information; Washington State Department of Health; West Virginia Health Care Authority; Wisconsin Department of Health and Family Services.

provider (see Figures 1.1 and 1.2 in Appendix C). For the condition indicators, the reference population rate is the same among all providers (see Figures 1.3 and 1.4).

Equal weight. In this case, each component indicator is assigned an identical weight based on the number of indicators. That is, the weight equals 1 divided by the number of indicators in the composite (e.g., 1/8 = 0.1250).

Numerator weight. A numerator weight is based on the relative frequency of the numerator for each component indicator in the reference population. In general, a numerator weight reflects the amount of harm in the outcome of interest, in this case mortality. For other types of outcomes the harm might reflect the amount of excess mortality or complications associated with the adverse event.

Denominator weight. A denominator weight is based on the relative frequency of the denominator for each component indicator in the reference population. In general, a denominator weight reflects the amount of risk of experiencing the outcome of interest in a given population. For example, the denominator weight might be based on the demographic composition of a health plan, the employees of a purchaser, a State, an individual hospital, or a single patient.

Factor weight. A factor weight is based on some sort of analysis that assigns each component indicator a weight that reflects the contribution of that indicator to the common variation among the indicators. The component indicator that is most predictive of that common variation is assigned the highest weight. The weights in Table 6 are based on a principal components factor analysis of the reliability-adjusted ratios.

Step 5. Construct the composite measure

The composite measure is the weighted average of the component indicators using the selected weights and the scaled and reliability-adjusted indicators.

> Composite = [indicator1 RAR × weight1] + [indicator2 RAR × weight2] + . . . + [indicatorN RAR × weightN]

The confidence interval of the composite is based on the standard error of the composite, which is the square root of the variance. The variance is computed based on the signal variance-covariance matrix and the reliability weights. Details of the computation are provided in Appendix D.

4.4. Sample Computation of the Composite Measure

This example demonstrates the construction of the composite for a representative provider beginning with the risk-adjusted rate and standard error for each IQI. An important consideration in the development of the composite measure methodology was that the computation of the composite and the weights be transparent and that a provider be able to trace the computation from the component indicators to the composite and back again.

Step 1. Compute the risk-adjusted rate and standard error

IQI	Average Annual Denominator	Observed Rate	Risk-Adjusted Rate	Rate Std. Error
IQI #15 In-Hosp Mort AMI	50	225.2	179.8	18.5
IQI #16 In-Hosp Mort CHF	334	83.9	105.2	7.1
IQI #17 In-Hosp Mort Stroke	103	132.3	198.9	19.8
IQI #18 In-Hosp Mort GI Hemorrhage	139	50.5	58.2	8.7
IQI #19 In-Hosp Mort Hip Fracture	93	46.4	76.3	13.2
IQI #20 In-Hosp Mort Pneumonia	506	106.0	147.1	7.9

Note: Observed and risk-adjusted rate are per 1,000.

This is the output a user would obtain from applying the AHRQ QI software (SAS and Windows) to the user's data.

Step 2. Scale the risk-adjusted rate using the reference population

IQI	Reference Population Rate	Risk-Adjusted Ratio	Ratio Std. Error
IQI #15 In-Hosp Mort AMI	93.5	1.924	0.198
IQI #16 In-Hosp Mort CHF	45.7	2.302	0.155
IQI #17 In-Hosp Mort Stroke	113.3	1.755	0.174
IQI #18 In-Hosp Mort GI Hemorrhage	31.6	1.842	0.276
IQI #19 In-Hosp Mort Hip Fracture	32.1	2.374	0.412
IQI #20 In-Hosp Mort Pneumonia	83.1	1.771	0.095

Step 3. Compute the reliability-adjusted ratio

Step S3A. Compute the reliability weight

IQI	Ratio Std. Error	Noise Variance	Signal Variance	Reliability Weight
IQI #15 In-Hosp Mort AMI	0.198	0.0394	0.0241	0.3797
IQI #16 In-Hosp Mort CHF	0.155	0.0239	0.0630	0.7251
IQI #17 In-Hosp Mort Stroke	0.174	0.0304	0.0474	0.6088
IQI #18 In-Hosp Mort GI Hemorrhage	0.276	0.0762	0.0508	0.4000
IQI #19 In-Hosp Mort Hip Fracture	0.412	0.1697	0.1083	0.3895
IQI #20 In-Hosp Mort Pneumonia	0.095	0.0090	0.0515	0.8517

Note: Noise variance is standard error squared (for details on calculating the noise variance, see Appendix D); reliability weight is signal variance/(signal variance + noise variance).

The noise variance is computed from the user's data as the square of the standard error. The signal variance is a reference population parameter that reflects the amount of provider-level variation remaining after the noise variance is removed. Note that the noise variance will vary by provider and by indicator.

Step S3B. Compute the reliability-adjusted ratio

IQI	Reliability Weight	Risk-Adjusted Ratio	Reference Population Ratio	Reliability-Adjusted Ratio
IQI #15 In-Hosp Mort AMI	0.3797	1.924	0.996	1.348
IQI #16 In-Hosp Mort CHF	0.7251	2.302	1.016	1.948
IQI #17 In-Hosp Mort Stroke	0.6088	1.755	1.005	1.462
IQI #18 In-Hosp Mort GI Hemorrhage	0.4000	1.842	1.008	1.342
IQI #19 In-Hosp Mort Hip Fracture	0.3895	2.374	1.036	1.557
IQI #20 In-Hosp Mort Pneumonia	0.8517	1.771	1.012	1.658

Note: Reliability-adjusted ratio is [risk-adjusted ratio × weight] + [reference population ratio × (1 – weight)].

The first "composite" is the weighted average of the provider's risk-adjusted ratio and the reference population ratio, where the weight reflects the reliability of the provider's risk-adjusted ratio. This "composite" is the reliability-adjusted ratio.

Step 4. Select the component weights

The weights are selected depending on the intended use of the composite. In this example, we use the numerator weight.

IQI	Numerator Weight
IQI #15 In-Hosp Mort AMI	0.1933
IQI #16 In-Hosp Mort CHF	0.1683
IQI #17 In-Hosp Mort Stroke	0.2036
IQI #18 In-Hosp Mort GI Hemorrhage	0.0525
IQI #19 In-Hosp Mort Hip Fracture	0.0320
IQI #20 In-Hosp Mort Pneumonia	0.3504

Step 5. Construct the composite measure

IQI	Numerator Weight (A)	Reliability-Adjusted Ratio (B)	(A) × (B)
IQI #15 In-Hosp Mort AMI	0.1933	1.348	0.261
IQI #16 In-Hosp Mort CHF	0.1683	1.948	0.328
IQI #17 In-Hosp Mort Stroke	0.2036	1.462	0.298
IQI #18 In-Hosp Mort GI Hemorrhage	0.0525	1.342	0.070
IQI #19 In-Hosp Mort Hip Fracture	0.0320	1.557	0.050
IQI #20 In-Hosp Mort Pneumonia	0.3504	1.658	0.581
Mortality for Selected Conditions			*1.588*
Standard Error			*0.066*
Confidence Interval at p<0.05		*1.458*	*1.717*

Note: For details on calculating the composite variance (standard error), see Appendix D.

The final composite is the weighted average of the component indicators, which is the sum of A × B for each indicator. Note the potential application of the composite construction for use in quality improvement. The final computation shows that pneumonia mortality is the largest single contributor to the composite both because the indicator was heavily weighted and because the performance of the provider was worse than average. The incentive created in using the composite is to allocate resources to reducing pneumonia mortality as the best mechanism to lower the composite score.

5. Performance of the AHRQ QI Composite Measures

5.1. Evaluation Criteria

Tables 7-9 in Appendix B and Figures 2.1-2.10 and 3.1-3.10 in Appendix C show the performance of each composite measure. The composite measures are evaluated using three criteria: discrimination, forecasting, and construct validity.

Discrimination is the ability of the composite measure to differentiate performance as measured by statistically significant deviations from the average performance.

Forecasting is the ability of the composite measure to predict performance for each of the component indicators. Ideally, the forecasting performance would reflect the weighting of the components, in the sense that forecasting would maximize the differences for the most highly weighted components.

Construct validity is the degree of association between the composite and other aggregate measures of quality. In this report we look primarily at the consistency in the composites with one another. A broader analysis of construct validity would examine the relationship between the composites and external measures of quality or other factors that might influence quality.

5.2. Results

Table 7 shows the discrimination performance of the two composite measures: *Mortality for Selected Procedures* and *Mortality for Selected Conditions*. The columns show the percentage of providers that are worse than average, average, or better than average based on the confidence interval for the composite measure. The discrimination performance varies depending on the weight used. For the procedure composite, the single indicator weight discriminates well because there is a strong volume-outcome relationship and because there are many hospitals performing only a few procedures. For the condition composite, the single indicator weight has the least ability to discriminate. The single indicator used as an example is “in-hospital mortality for pancreatic resection” for the procedure composite and “in-hospital mortality for hip fracture” for the condition composite. The numerator weight tends to have the greatest ability to discriminate, followed by the denominator weight or factor weight. Equal weight tends to discriminate the least among the alternative approaches.

The *Mortality for Selected Procedures* composite measure has slightly more variability in the ability to discriminate performance among the alternative approaches, and less ability to discriminate overall, than the *Mortality for Selected Conditions* composite measure. In general, however, both composites identify a large number of providers with performance that is better or worse than average. Figures 2.1-2.10 show the range of values for each composite for 400 randomly selected hospitals, with the 95 percent confidence interval, which illustrates the precision of the composites. Figures 3.1-3.10 show the distribution of alternative composites.

Table 8 shows the forecasting performance of the two composite measures. In this analysis each provider is assigned to a quintile (Q1-Q5) based on the performance on the composite in 2001-2003. The columns show the relative difference in the predicted risk-adjusted ratio in 2004 for the best and worst performing quintile relative to the middle 60 percent.

Forecasting performance varies depending on the weights used to construct the composite. In general, the composite is better at forecasting performance on component indicators that are more heavily weighted. In this sense the weights reflect the goals of the composite; more weight is assigned to component indicators where the goal is to reduce variability in performance.

Table 9 shows the correlation among the composite measures using the alternative weights. For *Mortality for Selected Procedures*, the correlations range from 0.196 to 0.899, varying by 0.703. For *Mortality for Selected Conditions*, the correlations tend to be very high regardless of the weight. Thus, the performance of individual hospitals on the composite tends to be highly correlated.

6. Concluding Comments

The intent of the AHRQ QI Composite Measure project was to develop a general methodology that could be used primarily to monitor performance in national and regional reporting, but that also could be applied to comparative reporting and quality improvement at the provider level. An important caveat in using the composite measures is that the measures are not intended to reflect any broader construct of quality than is reflected in the component indicators themselves. The composites are only as useful and valid as are the component indicators that make up the composite. The AHRQ QIs are currently undergoing review through the National Quality Forum (NQF) consensus development processes, and a number of validation studies of the component indicators are underway. The actual content of the composite (i.e., what component indicators to include) and the potential uses of the composite will depend on the results of that process for the component indicators.

As the AHRQ QIs and the data upon which they are based continue to improve, the composite measures will improve as potentially useful tools for decisionmaking in allocating quality improvement resources. For example, potential extensions of the composite measure method include the incorporation of process measures (from other data sources) and measures of cost (estimated from HCUP). We encourage AHRQ QI users to continue to submit comments and suggestions for improvement on the composite measures and the component indicators to the AHRQ QI support team at support@qualityindicators.ahrq.gov.

Appendix A. AHRQ QI Composite Measure Workgroup

Workgroup Members

- John Birkmeyer, University of Michigan
- Bruce Boissonnault, Niagara Health Quality Coalition
- John Bott, Employer Health Care Alliance Cooperative
- Dale Bratzler, Oklahoma Foundation for Medical Quality
- Sharon Cheng, Medicare Payment Advisory Commission (MedPAC)
- Elizabeth Clough, Wisconsin Collaborative for Healthcare Quality
- Nancy Dunton, University of Kansas Medical Center, School of Nursing
- John Hoerner, Hospital Industry Data Institute
- David Hopkins, Pacific Business Group on Health
- Gregg Meyer, Massachusetts General Physicians Organization
- Elizabeth Mort, Massachusetts General
- Janet Muri, National Perinatal Information Center
- Vi Naylor, Georgia Hospital Association
- Eric Peterson, Duke University Medical Center
- Martha Radford, New York University Hospitals Center
- Gulzar Shah, National Association of Health Data Organizations
- Paul Turner, Vermont Program for Quality in Health Care

Liaison Members

- Justine Carr, National Committee on Vital and Health Statistics
- Robert Hungate, National Committee on Vital and Health Statistics
- Sheila Roman, Centers for Medicare & Medicaid Services
- Amy Rosen, Bedford Veterans Affairs Medical Center
- Stephen Schmaltz, Joint Commission on Accreditation of Healthcare Organizations
- Jane Sisk, National Center for Health Statistics
- Ernie Moy, Agency for Healthcare Research and Quality

Technical Advisors

- John Adams, RAND Corporation
- Bob Houchens, Medstat
- Bill Rogers, Rogers Associates
- Chunliu Zhan, Agency for Healthcare Research and Quality

AHRQ QI Support

- Mamatha Pancholi, AHRQ QI Project Officer
- Marybeth Farquhar, AHRQ NQF Project Officer
- Jeffrey Geppert, Project Director, Battelle Memorial Institute
- Theresa Schaaf, Project Manager, Battelle Memorial Institute
- Douglas O. Staiger, Technical Consultant, Dartmouth College

Appendix B. IQI Composite Tables

Table 1. Reference Population

IQI	Numerator	Denominator	Rate
IQI #08 In-Hosp Mort Esophageal Resection	344	4,191	82.08
IQI #09 In-Hosp Mort Pancreatic Resection	766	12,053	63.55
IQI #11 In-Hosp Mort AAA Repair	8,094	101,120	80.04
IQI #12 In-Hosp Mort CABG	29,460	857,107	34.37
IQI #13 In-Hosp Mort Craniotomy	20,713	282,594	73.30
IQI #14 In-Hosp Mort Hip Replacement	1,323	466,738	2.83
IQI #30 In-Hosp Mort PTCA	24,658	1,804,021	13.67
IQI #31 In-Hosp Mort Carotid Endarterectomy	2,458	355,596	6.91
IQI #15 In-Hosp Mort AMI	144,761	1,548,304	93.50
IQI #16 In-Hosp Mort CHF	126,036	2,758,388	45.69
IQI #17 In-Hosp Mort Stroke	152,492	1,345,994	113.29
IQI #18 In-Hosp Mort GI Hemorrhage	39,286	1,243,668	31.59
IQI #19 In-Hosp Mort Hip Fracture	23,977	745,896	32.15
IQI #20 In-Hosp Mort Pneumonia	262,398	3,158,119	83.09

Source: HCUP State Inpatient Databases, 2001-2003; rate per 1,000.

Table 2. Provider-Level Rates

		Risk Adjusted		Reliability Adjusted	
IQI	**Hospitals**	**Rate**	**Std. Dev.**	**Rate**	**Std. Dev.**
IQI #08 In-Hosp Mort Esophageal Resection	424	99.44	198.00	97.02	17.53
IQI #09 In-Hosp Mort Pancreatic Resection	857	77.56	168.36	77.87	13.53
IQI #11 In-Hosp Mort AAA Repair	1,963	91.01	103.64	90.23	13.71
IQI #12 In-Hosp Mort CABG	1,021	36.77	18.46	37.04	8.47
IQI #13 In-Hosp Mort Craniotomy	1,551	73.32	69.77	72.17	12.15
IQI #14 In-Hosp Mort Hip Replacement	3,064	3.48	19.76	2.77	1.99
IQI #30 In-Hosp Mort PTCA	1,261	14.15	10.87	14.87	2.65
IQI #31 In-Hosp Mort Carotid Endarterectomy	2,339	9.61	66.14	7.79	1.12
IQI #15 In-Hosp Mort AMI	3,962	108.88	65.77	95.08	10.66
IQI #16 In-Hosp Mort CHF	4,369	55.24	58.29	47.45	10.29
IQI #17 In-Hosp Mort Stroke	4,230	130.01	100.19	115.66	20.93
IQI #18 In-Hosp Mort GI Hemorrhage	4,162	33.44	41.61	31.90	3.80
IQI #19 In-Hosp Mort Hip Fracture	3,583	35.46	46.15	33.27	6.82
IQI #20 In-Hosp Mort Pneumonia	4,463	85.57	52.98	84.60	17.63

Source: HCUP State Inpatient Databases, 2001-2003; rate per 1,000.

Table 3. Provider-Level Correlation for Procedures

IQI	IQI #08	IQI #09	IQI #11	IQI #12	IQI #13	IQI #14	IQI #30	IQI #31
IQI #08 In-Hosp Mort Esophageal Resection	1.000	0.250	0.092	0.182	0.056	0.067	-0.005	0.117
IQI #09 In-Hosp Mort Pancreatic Resection		1.000	0.022	0.013	0.103	0.021	-0.030	0.076
IQI #11 In-Hosp Mort AAA Repair			1.000	0.131	0.105	0.012	0.060	0.046
IQI #12 In-Hosp Mort CABG				1.000	0.171	0.067	0.343	0.190
IQI #13 In-Hosp Mort Craniotomy					1.000	0.066	0.159	0.050
IQI #14 In-Hosp Mort Hip Replacement						1.000	0.077	0.037
IQI #30 In-Hosp Mort PTCA							1.000	0.094
IQI #31 In-Hosp Mort Carotid Endarterectomy								1.000

Source: HCUP State Inpatient Databases, 2001-2003.

Table 4. Provider-Level Correlation for Conditions

IQI	IQI #15	IQI #16	IQI #17	IQI #18	IQI #19	IQI #20
IQI #15 In-Hosp Mort AMI	1.000	0.460	0.378	0.288	0.227	0.450
IQI #16 In-Hosp Mort CHF		1.000	0.473	0.401	0.273	0.637
IQI #17 In-Hosp Mort Stroke			1.000	0.276	0.263	0.516
IQI #18 In-Hosp Mort GI Hemorrhage				1.000	0.170	0.408
IQI #19 In-Hosp Mort Hip Fracture					1.000	0.288
IQI #20 In-Hosp Mort Pneumonia						1.000

Source: HCUP State Inpatient Databases, 2001-2003.

Table 5. Reliability Weight by Average Annual Denominator

Average Annual Denominator Size (by quartile)

IQI	Hospitals	Q1	Q2	Q3	Q4
IQI #08 In-Hosp Mort Esophageal Resection	424	1.0	1.4	2.4	8.4
IQI #09 In-Hosp Mort Pancreatic Resection	857	1.1	1.8	3.1	12.7
IQI #11 In-Hosp Mort AAA Repair	1,963	1.9	5.6	13.8	47.3
IQI #12 In-Hosp Mort CABG	1,021	51.8	148.8	271.5	645.8
IQI #13 In-Hosp Mort Craniotomy	1,551	3.7	16.2	40.6	182.3
IQI #14 In-Hosp Mort Hip Replacement	3,064	4.5	17.0	43.1	138.5
IQI #30 In-Hosp Mort PTCA	1,261	33.0	198.4	461.6	1,212.2
IQI #31 In-Hosp Mort Carotid Endarterectomy	2,339	5.1	19.5	47.4	130.6
IQI #15 In-Hosp Mort AMI	3,962	5.1	23.3	79.0	413.4
IQI #16 In-Hosp Mort CHF	4,369	15.6	70.3	203.1	552.4
IQI #17 In-Hosp Mort Stroke	4,230	7.7	33.8	97.8	284.9
IQI #18 In-Hosp Mort GI Hemorrhage	4,162	8.2	39.3	103.2	247.6
IQI #19 In-Hosp Mort Hip Fracture	3,583	8.2	34.8	74.3	160.2
IQI #20 In-Hosp Mort Pneumonia	4,463	25.2	105.8	254.1	558.3

Average Reliability Weight

IQI	Q1	Q2	Q3	Q4	Weighted Average
IQI #08 In-Hosp Mort Esophageal Resection	0.0306	0.0450	0.0623	0.1551	0.1565
IQI #09 In-Hosp Mort Pancreatic Resection	0.0230	0.0313	0.0496	0.1196	0.1421
IQI #11 In-Hosp Mort AAA Repair	0.0446	0.0942	0.1666	0.3096	0.2924
IQI #12 In-Hosp Mort CABG	0.2807	0.5502	0.6852	0.8235	0.7464
IQI #13 In-Hosp Mort Craniotomy	0.0862	0.2853	0.4818	0.7074	0.6915
IQI #14 In-Hosp Mort Hip Replacement	0.0589	0.1607	0.3120	0.5258	0.4767
IQI #30 In-Hosp Mort PTCA	0.1038	0.3710	0.5563	0.7220	0.6599
IQI #31 In-Hosp Mort Carotid Endarterectomy	0.0086	0.0315	0.0721	0.1732	0.1523
IQI #15 In-Hosp Mort AMI	0.0620	0.2412	0.5030	0.7374	0.7160
IQI #16 In-Hosp Mort CHF	0.1182	0.3692	0.6474	0.8334	0.7550
IQI #17 In-Hosp Mort Stroke	0.1102	0.3562	0.6458	0.8649	0.7800
IQI #18 In-Hosp Mort GI Hemorrhage	0.0386	0.1616	0.3535	0.5761	0.4878
IQI #19 In-Hosp Mort Hip Fracture	0.0968	0.2908	0.4758	0.6551	0.5619
IQI #20 In-Hosp Mort Pneumonia	0.2128	0.5509	0.7832	0.8983	0.8222

Source: HCUP State Inpatient Databases, 2001-2003.

Table 6. Alternative Composite Weights

IQI	Single Indicator Weight	Equal Weight	Numerator Weight	Denominator Weight	Factor Weight
IQI #08 In-Hosp Mort Esophageal Resection	0.0000	0.1250	0.0039	0.0011	0.1236
IQI #09 In-Hosp Mort Pancreatic Resection	**1.0000**	0.1250	0.0087	0.0031	0.1534
IQI #11 In-Hosp Mort AAA Repair	0.0000	0.1250	0.0922	0.0260	**0.1904**
IQI #12 In-Hosp Mort CABG	0.0000	0.1250	**0.3355**	0.2207	0.1679
IQI #13 In-Hosp Mort Craniotomy	0.0000	0.1250	0.2359	0.0728	0.0035
IQI #14 In-Hosp Mort Hip Replacement	0.0000	0.1250	0.0151	0.1202	0.0076
IQI #30 In-Hosp Mort PTCA	0.0000	0.1250	0.2808	**0.4645**	0.1786
IQI #31 In-Hosp Mort Carotid Endarterectomy	0.0000	0.1250	0.0280	0.0916	0.1751
IQI #15 In-Hosp Mort AMI	0.0000	0.1667	0.1933	0.1434	0.1598
IQI #16 In-Hosp Mort CHF	0.0000	0.1667	0.1683	0.2554	0.1828
IQI #17 In-Hosp Mort Stroke	0.0000	0.1667	0.2036	0.1246	0.1764
IQI #18 In-Hosp Mort GI Hemorrhage	0.0000	0.1667	0.0525	0.1152	0.1475
IQI #19 In-Hosp Mort Hip Fracture	**1.0000**	0.1667	0.0320	0.0691	0.1444
IQI #20 In-Hosp Mort Pneumonia	0.0000	0.1667	**0.3504**	**0.2924**	**0.1891**

Source: HCUP State Inpatient Databases, 2001-2003. For each indicator, the most highly weighted component is in **bold**.

Table 7. Discrimination Performance of Alternative Composites

Composite	Providers	Better Than Average	Average	Worse Than Average
Mortality for Selected Procedures				
Single Indicator Weight	857	0.93%	69.08%	29.99%
Equal Weight	2,688	0.41%	97.62%	1.97%
Numerator Weight	1,978	4.25%	82.25%	13.50%
Denominator Weight	2,649	1.40%	88.00%	10.61%
Factor Weight	2,146	0.70%	78.56%	20.74%
Mortality for Selected Conditions				
Single Indicator Weight	3,583	2.60%	94.45%	2.96%
Equal Weight	4,384	9.01%	81.87%	9.12%
Numerator Weight	4,426	12.49%	73.50%	14.01%
Denominator Weight	4,415	11.26%	76.17%	12.57%
Factor Weight	4,392	9.93%	79.90%	10.18%

Source: HCUP State Inpatient Databases, 2001-2003.

Table 8A. Forecasting Performance of Alternative Composites

IQI	IQI #08	IQI #09	IQI #11	IQI #12	IQI #13	IQI #14	IQI #30	IQI #31
Mortality for Selected Procedures								
Single Indicator Weight								
Best 20%	-0.335*	**-0.519***	0.129*	0.099*	-0.181*	-0.049	0.089*	0.226*
Worst 20%	0.233*	**0.382***	0.134*	0.051*	0.050*	0.073	0.094*	0.038*
Equal Weight								
Best 20%	-0.250*	-0.285*	-0.140*	-0.177*	-0.093*	-0.344*	-0.144*	-0.084*
Worst 20%	0.054*	0.129*	0.161*	0.261*	0.082*	0.480*	0.218*	0.117*
Numerator Weight								
Best 20%	-0.043*	-0.069*	-0.096*	**-0.238***	-0.130*	-0.169*	-0.209*	-0.076*
Worst 20%	0.024	0.058*	0.091*	**0.351***	0.138*	0.157*	0.219*	0.069*
Denominator Weight								
Best 20%	-0.050*	-0.087*	-0.081*	-0.217*	-0.094*	-0.258*	**-0.233***	-0.087*
Worst 20%	0.019	0.043*	0.085*	0.298*	0.086*	0.391*	**0.264***	0.074*
Factor Weight								
Best 20%	-0.264*	-0.284*	**-0.142***	-0.165*	-0.073*	-0.219*	-0.134*	-0.079*
Worst 20%	0.057*	0.146*	**0.162***	0.305*	0.047*	0.057	0.235*	0.113*

Source: HCUP State Inpatient Databases, 2001-2003.
*Significant at p<.05. The forecast predicts performance in 2004 based on performance in 2001-2003 (by quintile) using five alternative measure composite weights. For each indicator, the most highly weighted component is in **bold**.

Table 8B. Forecasting Performance of Alternative Composites

IQI	IQI #15	IQI #16	IQI #17	IQI #18	IQI #19	IQI #20
Mortality for Selected Conditions						
Single Indicator Weight						
Best 20%	-0.079*	-0.128*	-0.176*	-0.074*	**-0.315***	-0.126*
Worst 20%	0.074*	0.152*	0.114*	0.078*	**0.348***	0.151*
Equal Weight						
Best 20%	-0.139*	-0.247*	-0.200*	-0.139*	-0.211*	-0.211*
Worst 20%	0.135*	0.259*	0.185*	0.147*	0.223*	0.273*
Numerator Weight						
Best 20%	-0.134*	-0.248*	-0.202*	-0.116*	-0.172*	**-0.236***
Worst 20%	0.132*	0.258*	0.188*	0.133*	0.164*	**0.300***
Denominator Weight						
Best 20%	-0.134*	-0.261*	-0.188*	-0.128*	-0.174*	**-0.233***
Worst 20%	0.127*	0.274*	0.174*	0.141*	0.185*	**0.294***
Factor Weight						
Best 20%	-0.138*	-0.251*	-0.204*	-0.132*	-0.203*	**-0.217***
Worst 20%	0.135*	0.262*	0.186*	0.147*	0.215*	**0.277***

Source: HCUP State Inpatient Databases, 2001-2003.
*Significant at p<.05. The forecast predicts performance in 2004 based on performance in 2001-2003 (by quintile) using five alternative measure composite weights. For each indicator, the most highly weighted component is in **bold**.

Table 9. Correlation of Alternative Composites

Composite	Single Indicator Weight	Equal Weight	Numerator Weight	Denominator Weight	Factor Weight
Mortality for Selected Procedures					
Single Indicator Weight	1.000	0.565	0.219	0.196	0.633
Equal Weight		1.000	0.744	0.831	0.853
Numerator Weight			1.000	0.899	0.809
Denominator Weight				1.000	0.754
Factor Weight					1.000
Mortality for Selected Conditions					
Single Indicator Weight	1.000	0.711	0.567	0.598	0.687
Equal Weight		1.000	0.971	0.981	0.999
Numerator Weight			1.000	0.992	0.980
Denominator Weight				1.000	0.988
Factor Weight					1.000

Source: HCUP State Inpatient Databases, 2001-2003.

Appendix C. IQI Composite Figures

1. *Single Indicator Composites*

Figure 1.1 - IQI #9 Pancreatic Resection Mortality

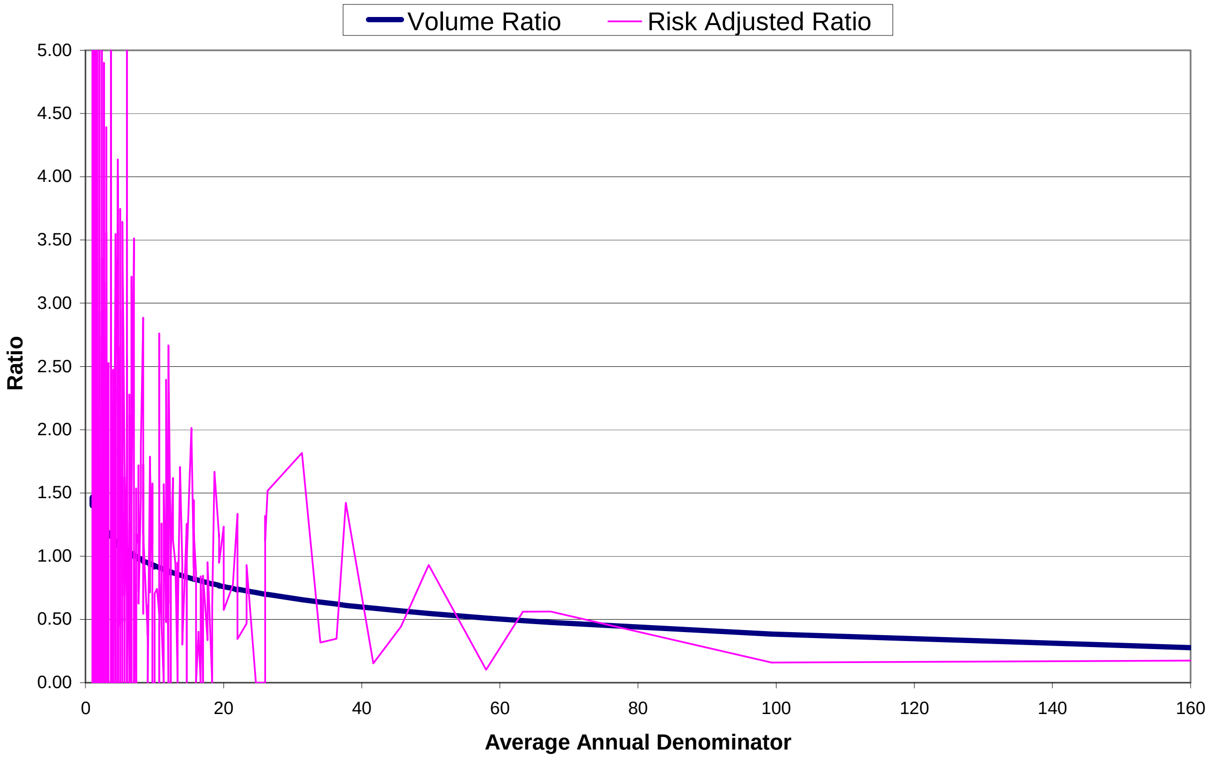


Figure 1.2 - IQI #9 Pancreatic Resection Mortality

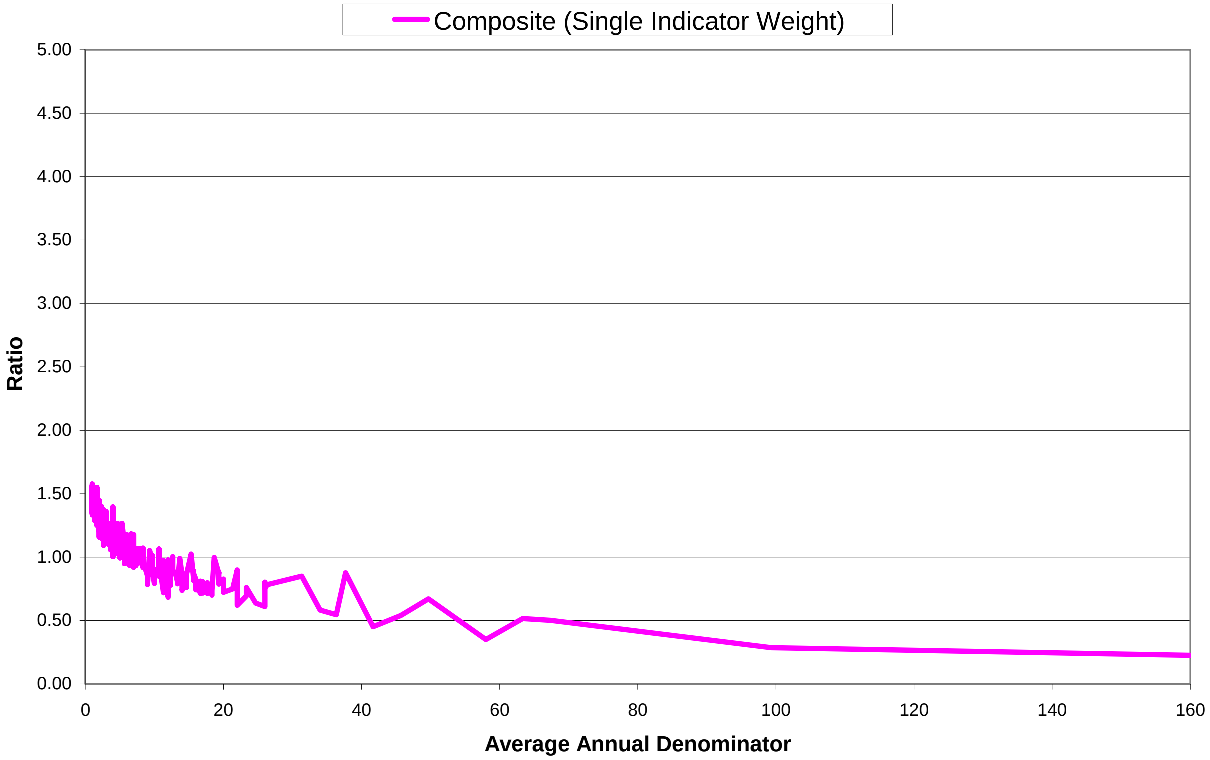


Figure 1.3 - IQI #19 Hip Fracture Mortality

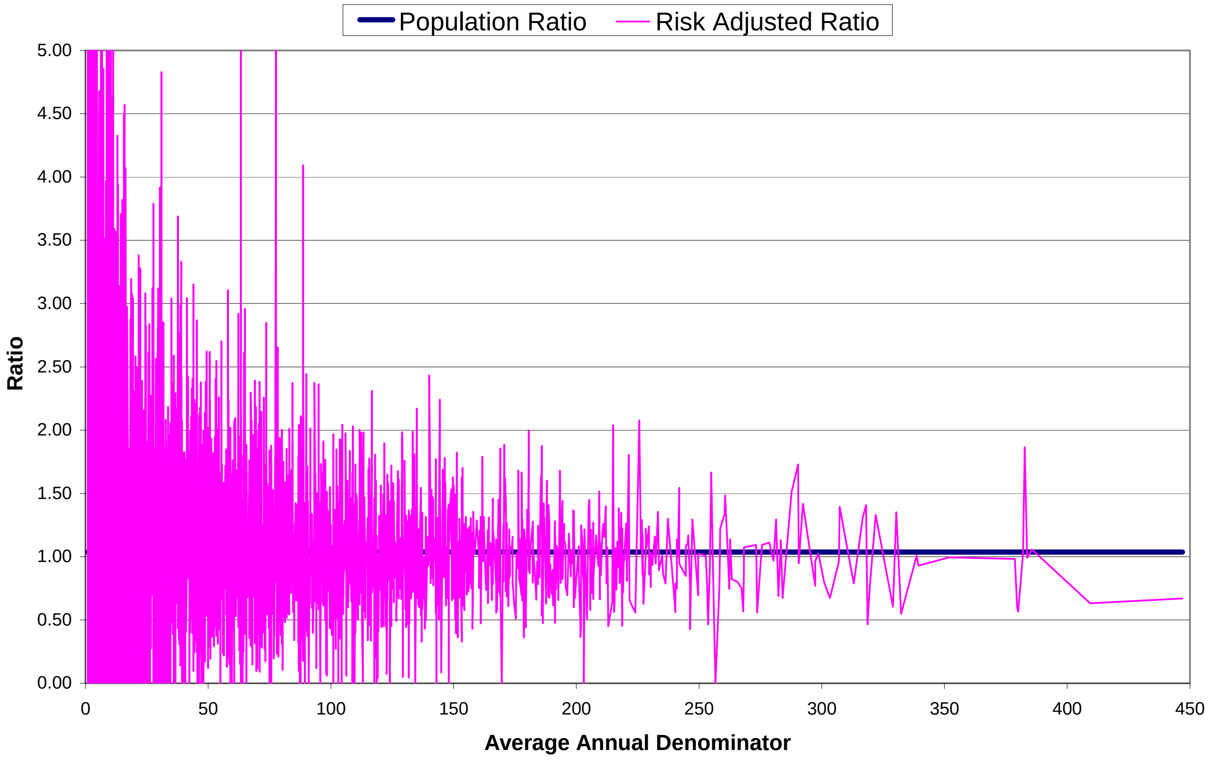
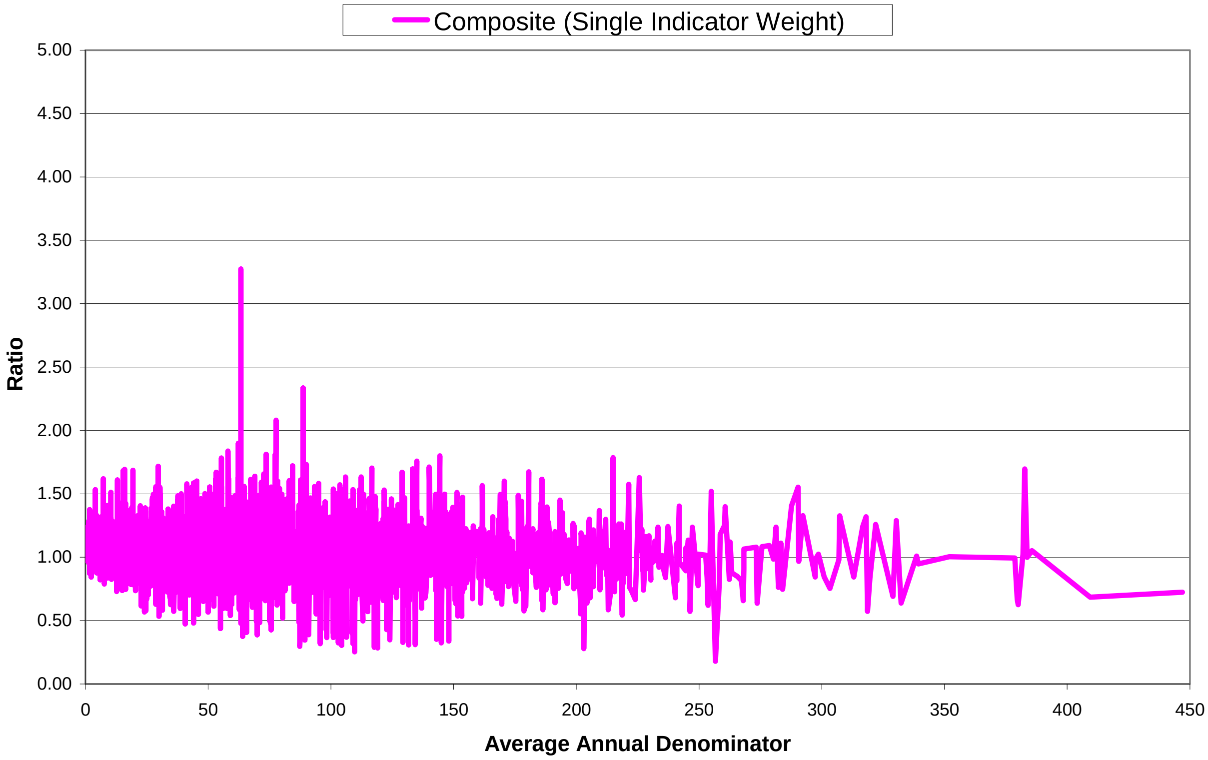


Figure 1.4 - IQI #19 Hip Fracture Mortality



2.. *Precision of Alternative Composites*

Figure 2.1 - Mortality for Selected Procedures, Single Indicator Weight

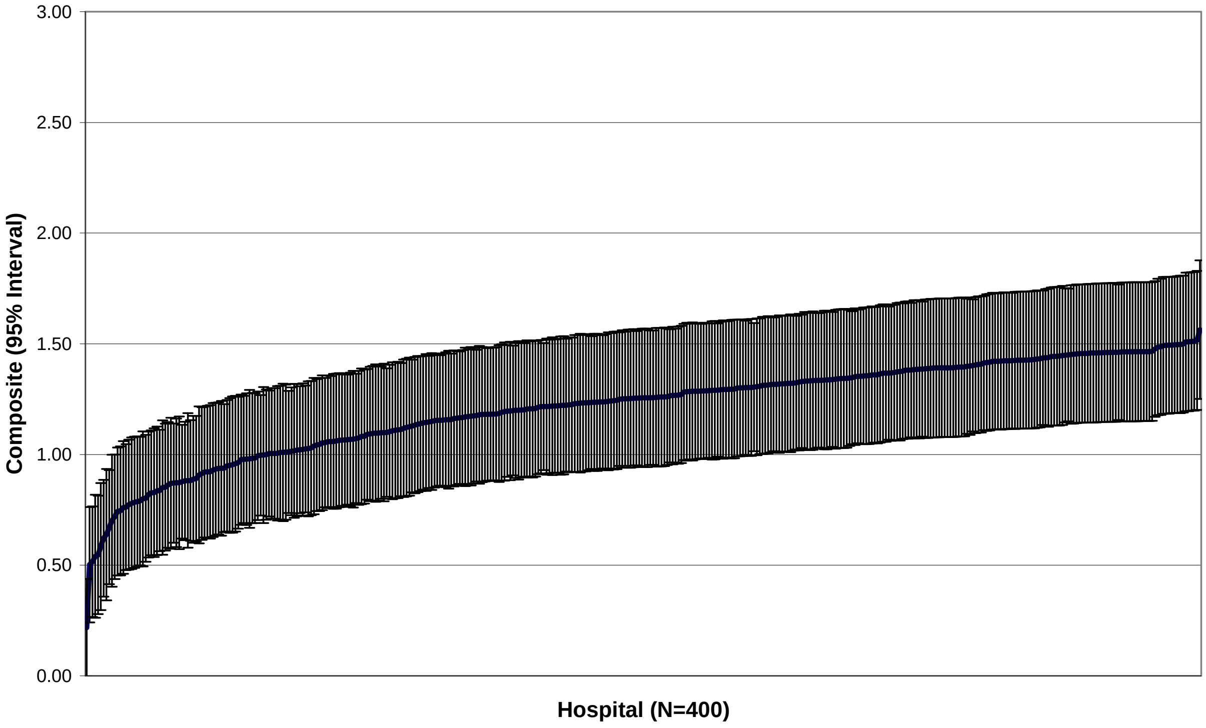


Figure 2.2 - Mortality for Selected Procedures, Equal Weight

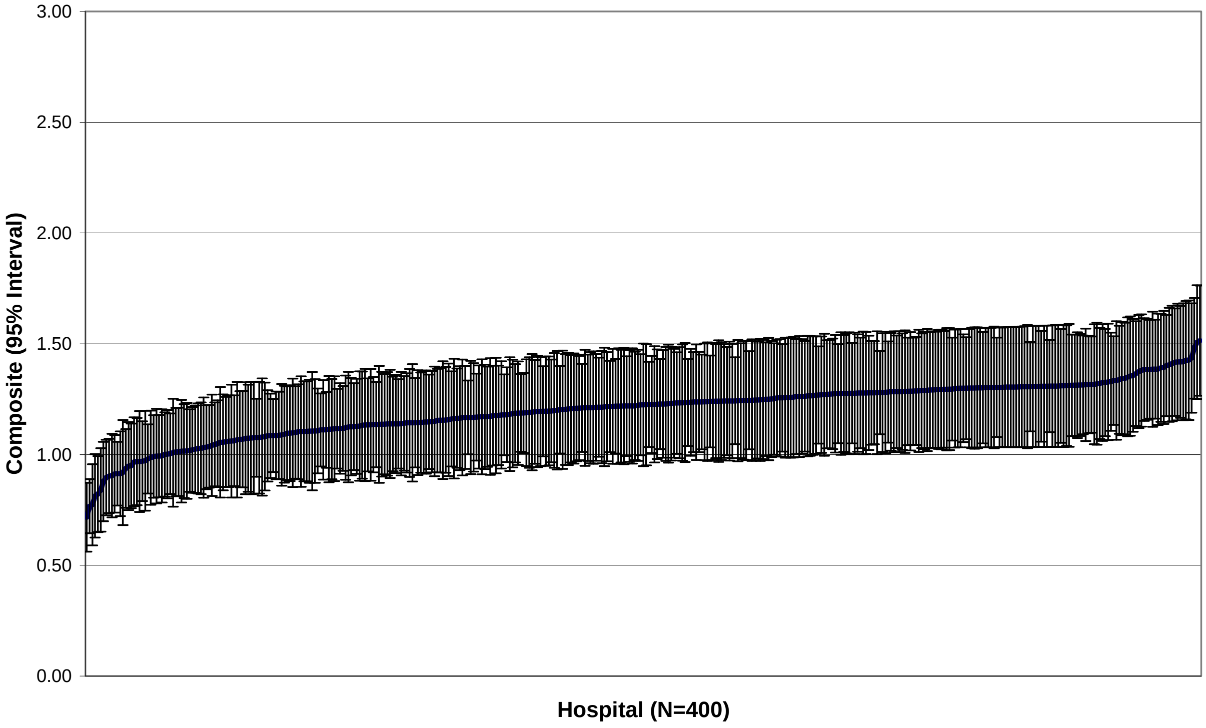


Figure 2.3 - Mortality for Selected Procedures, Numerator Weight

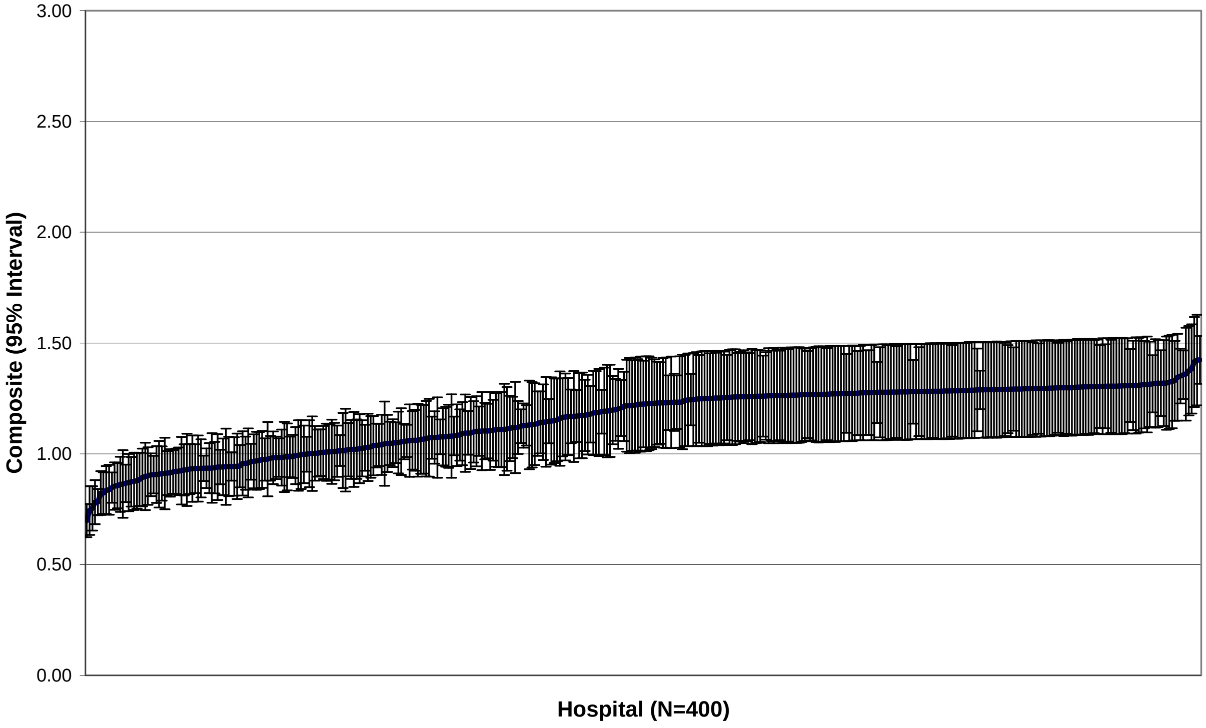


Figure 2.4 - Mortality for Selected Procedures, Denominator Weight

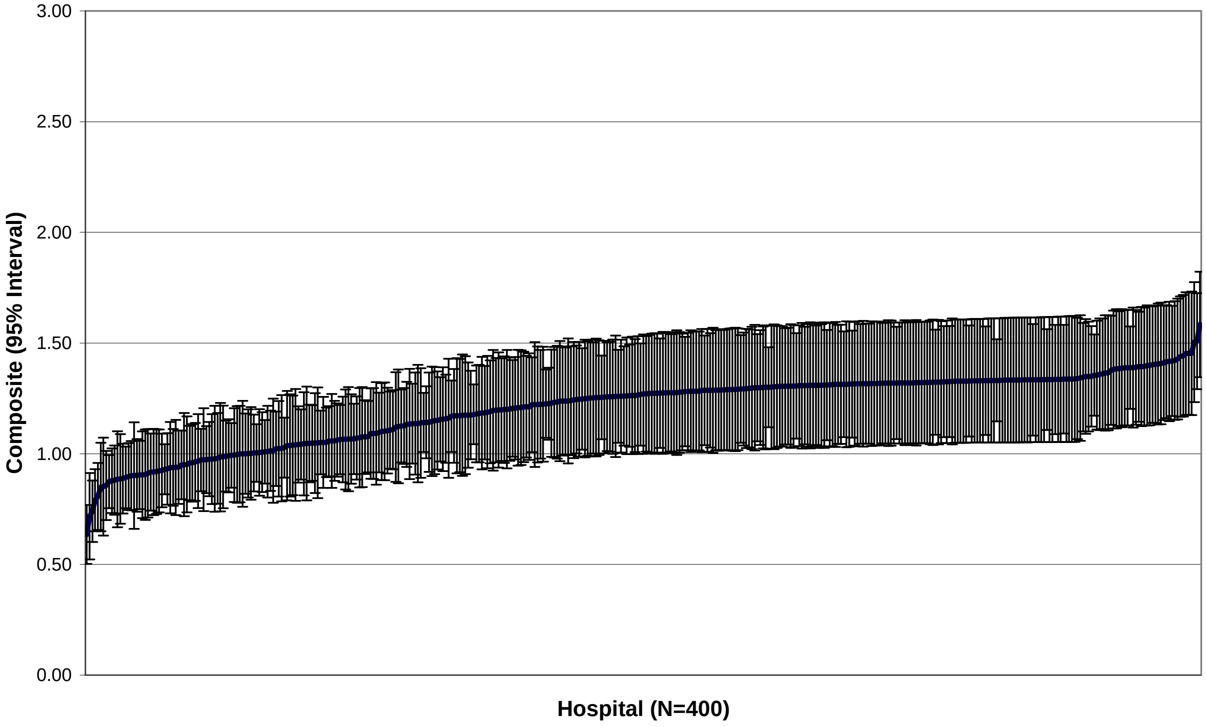


Figure 2.5 - Mortality for Selected Procedures, Factor Weight

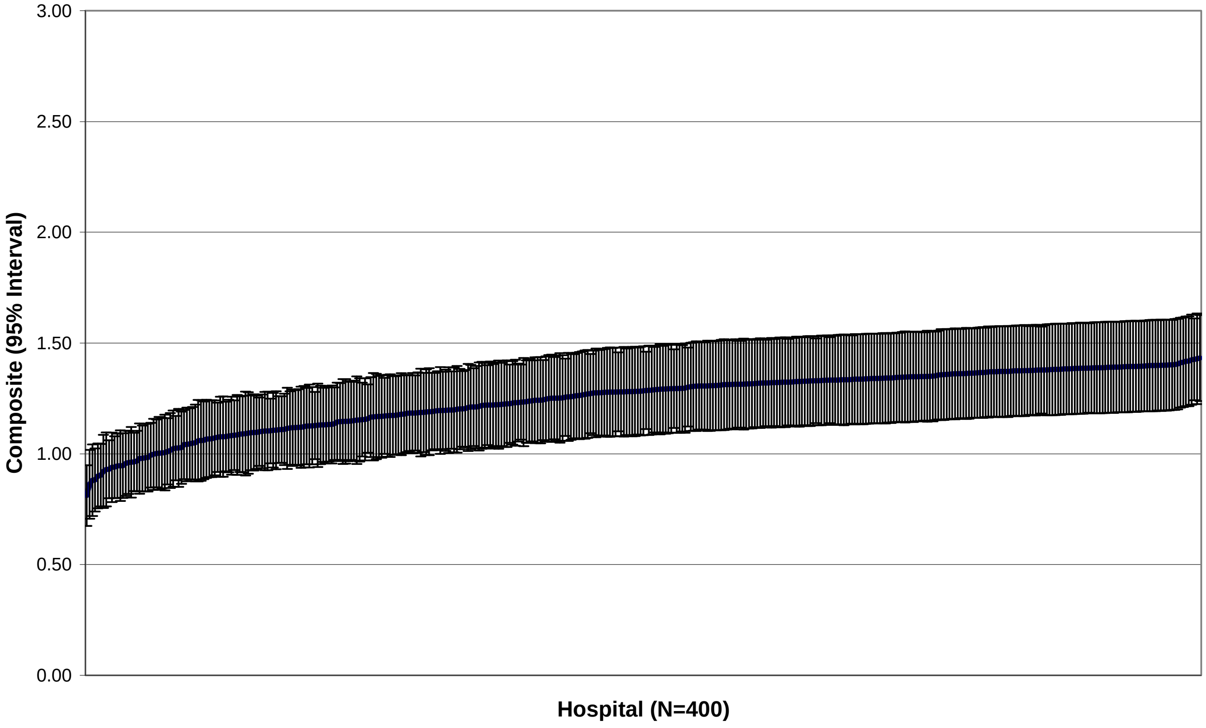


Figure 2.6 - Mortality for Selected Conditions, Single Indicator Weight

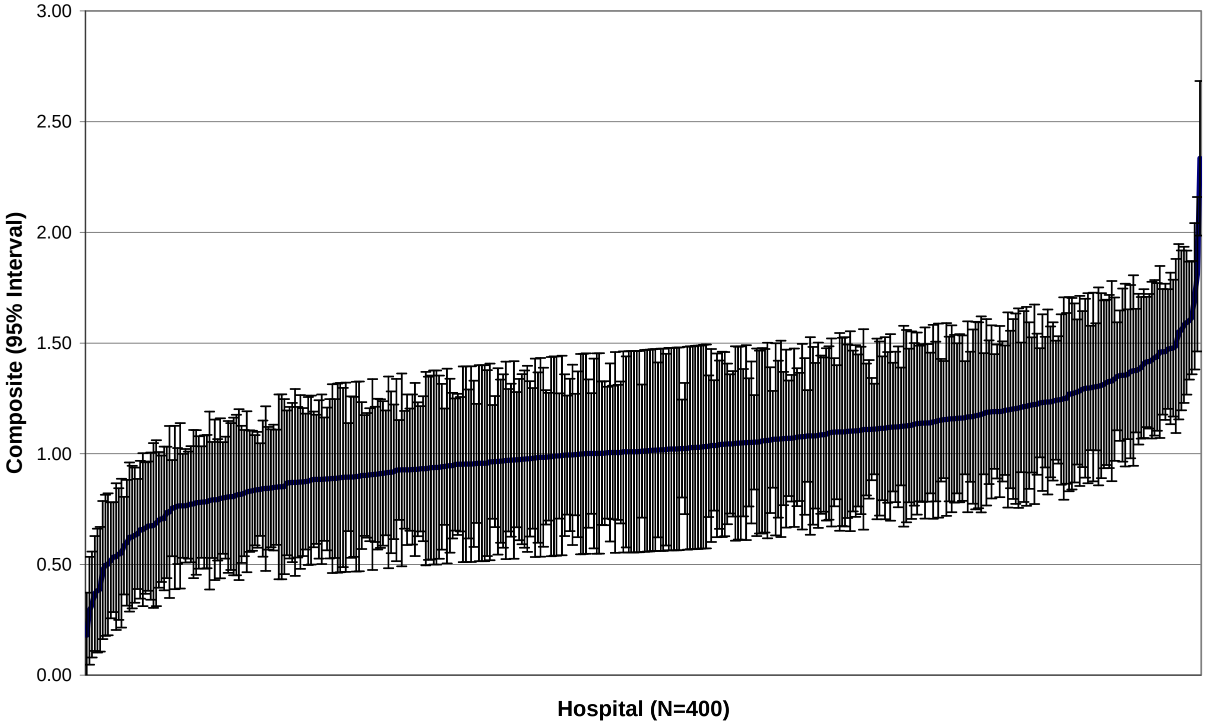


Figure 2.7 - Mortality for Selected Conditions, Equal Weight

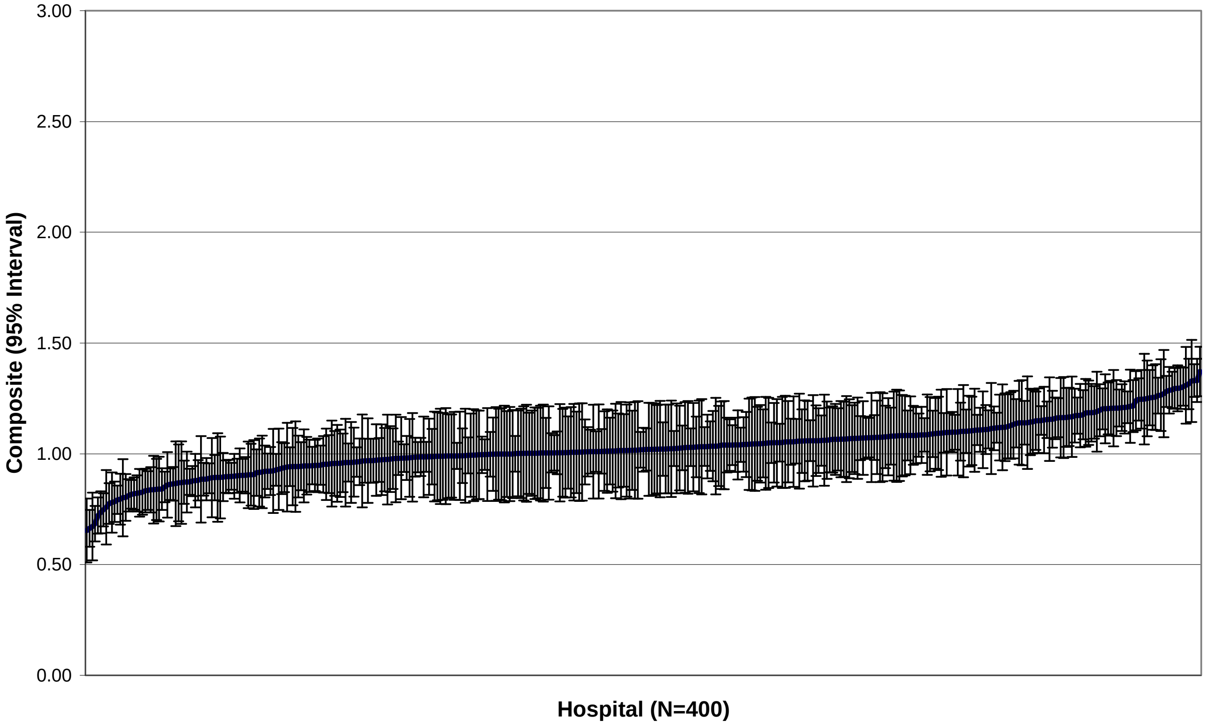


Figure 2.8 - Mortality for Selected Conditions, Numerator Weight

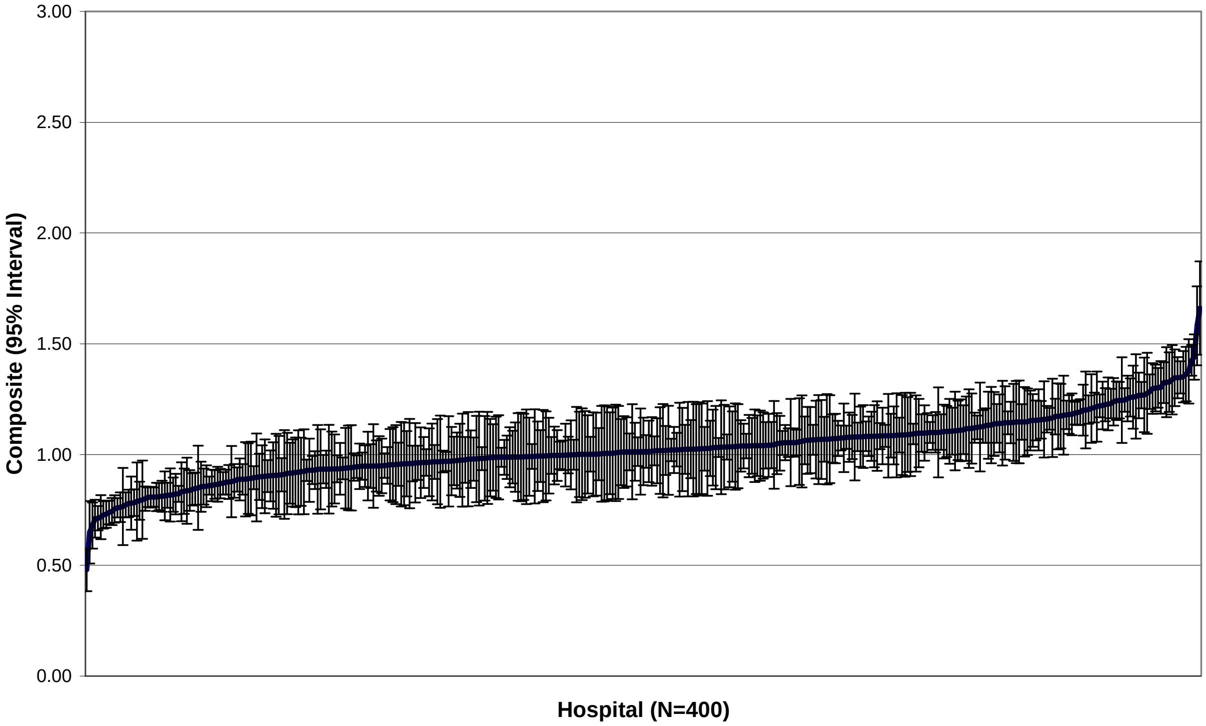


Figure 2.9 - Mortality for Selected Conditions, Denominator Weight

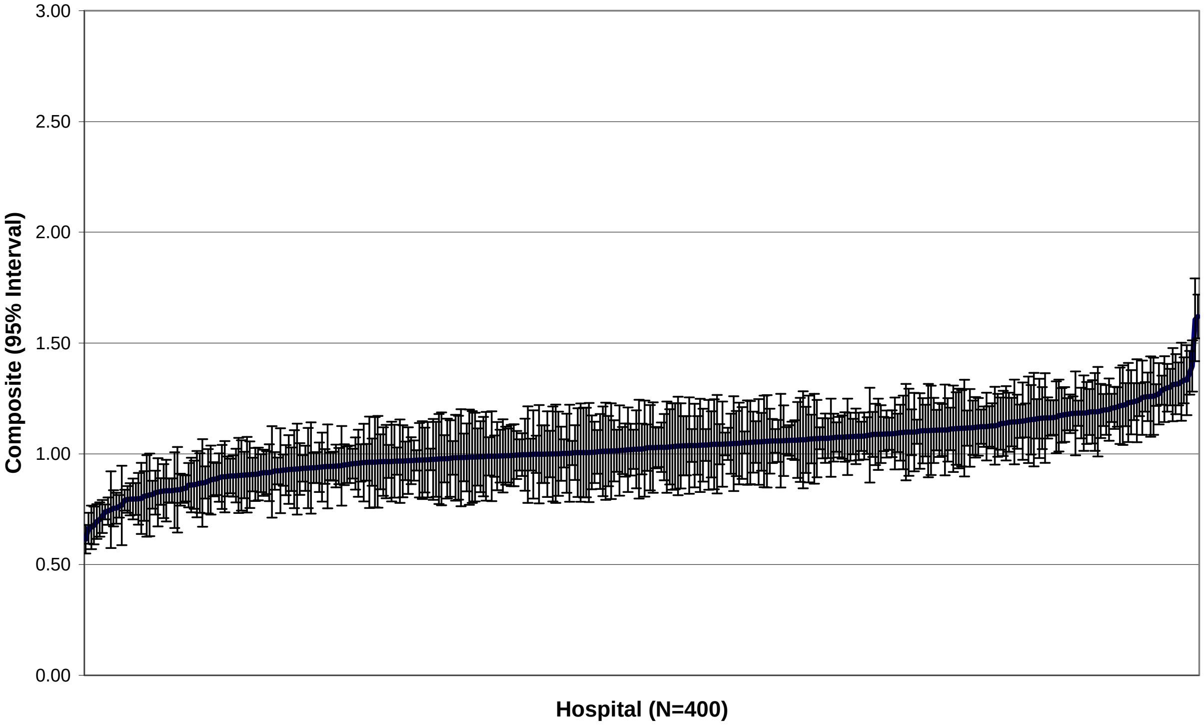
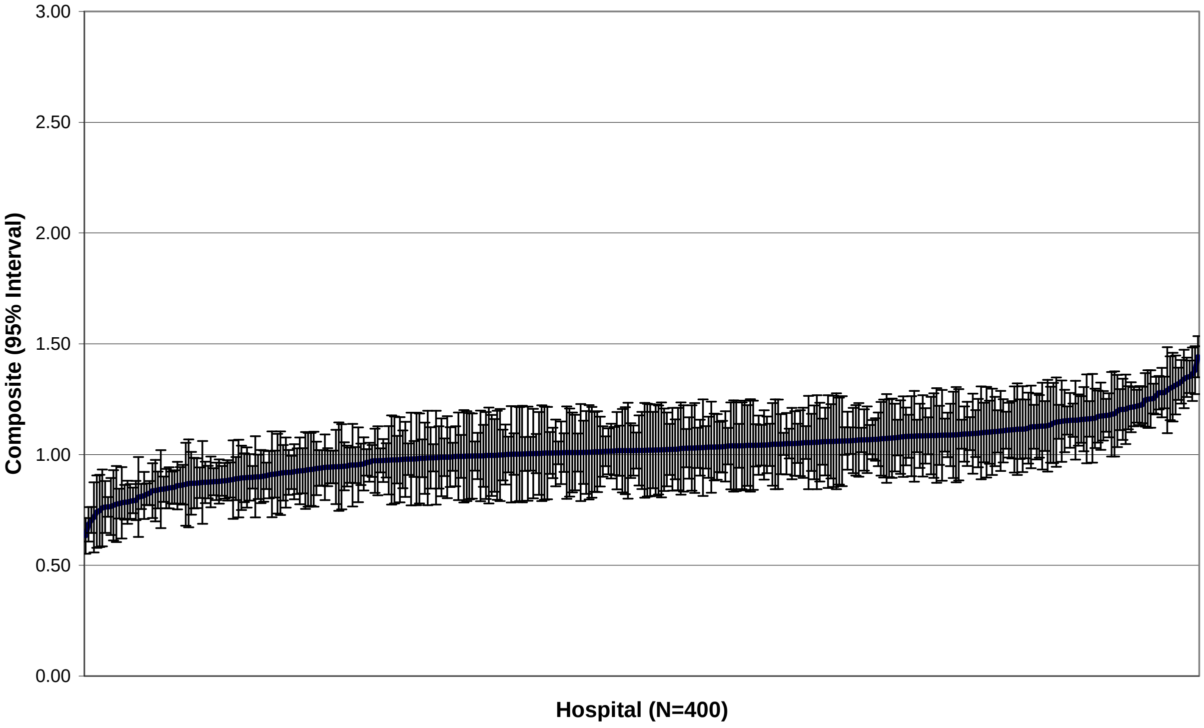


Figure 2.10 - Mortality for Selected Conditions, Factor Weight



3. Distribution of Alternative Composites

Figure 3.1 - Mortality for Selected Procedures, Single Indicator Weight

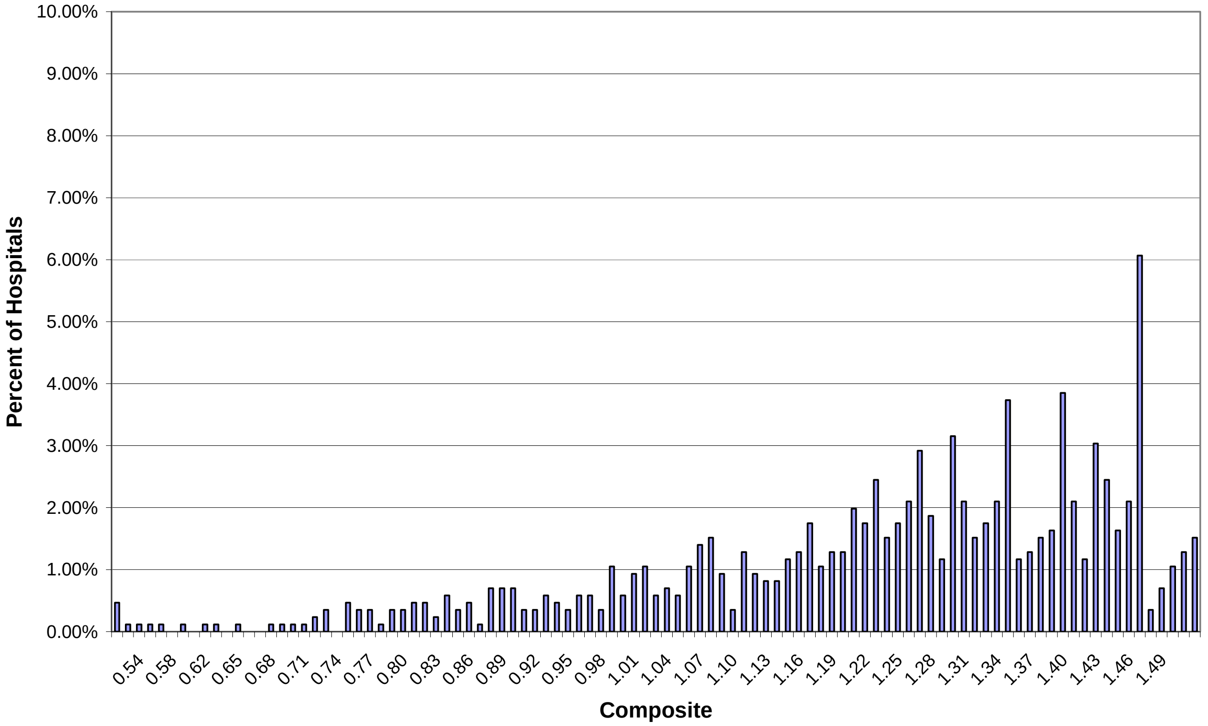


Figure 3.2 - Mortality for Selected Procedures, Equal Weight

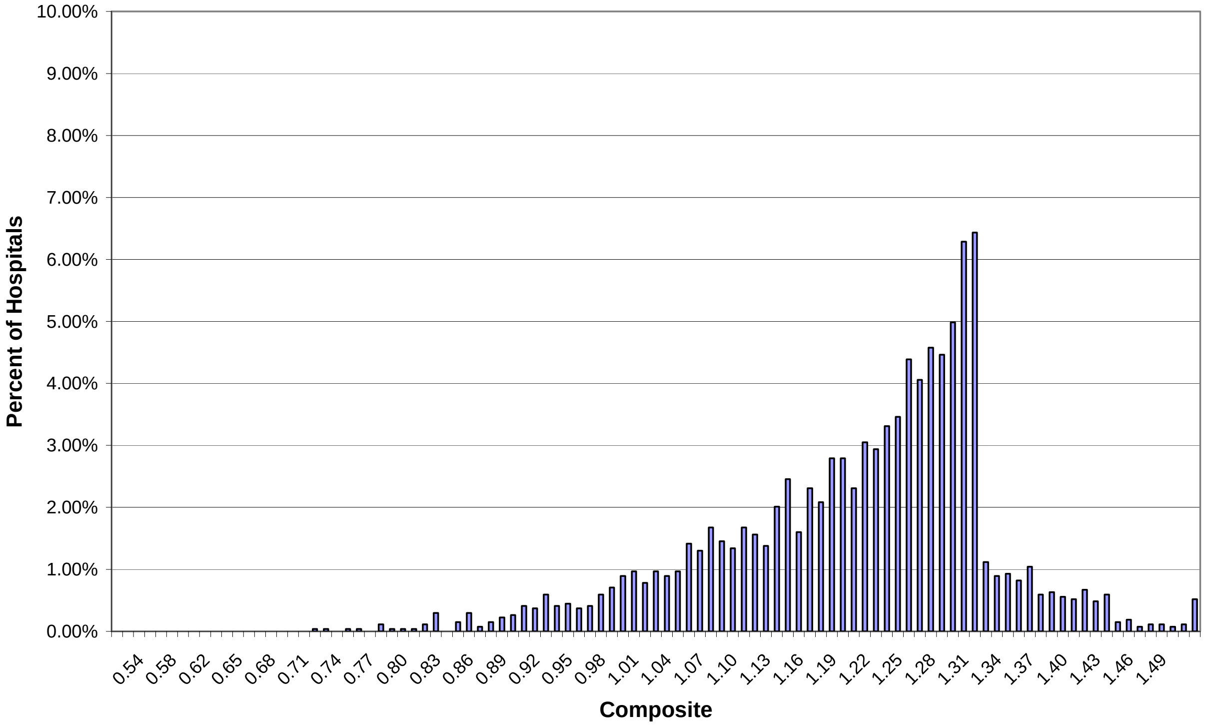


Figure 3.3 - Mortality for Selected Procedures, Numerator Weight

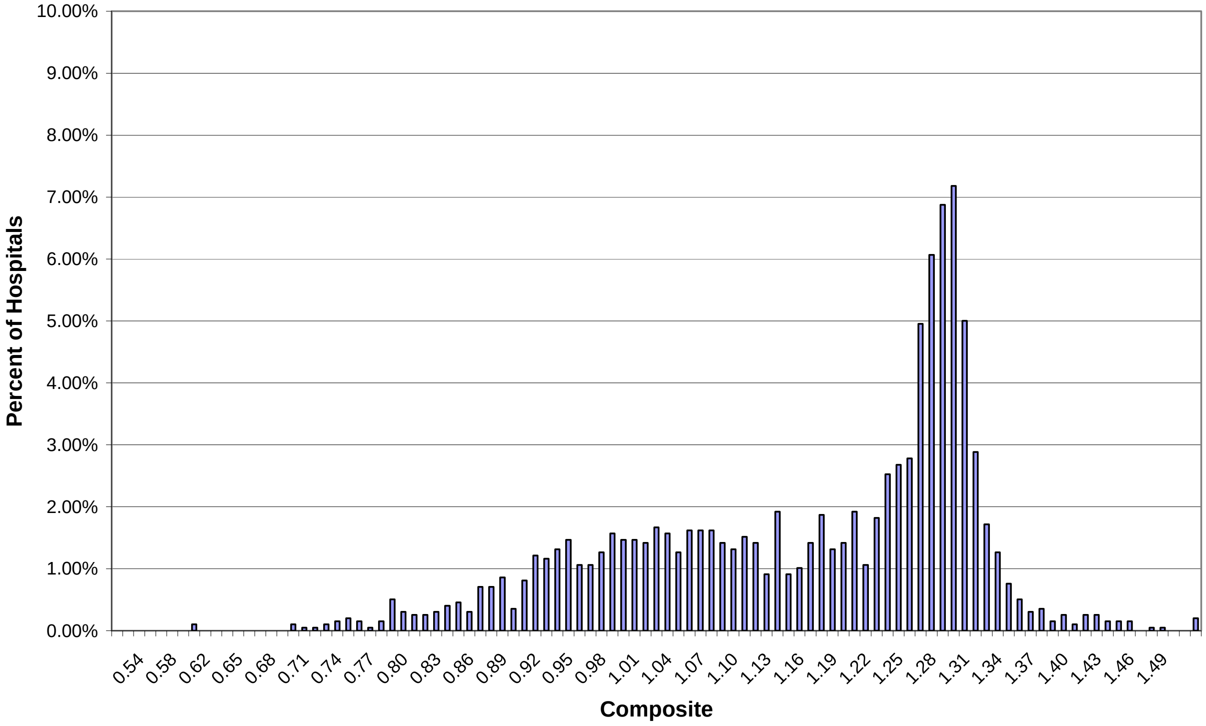


Figure 3.4 - Mortality for Selected Procedures, Denominator Weight

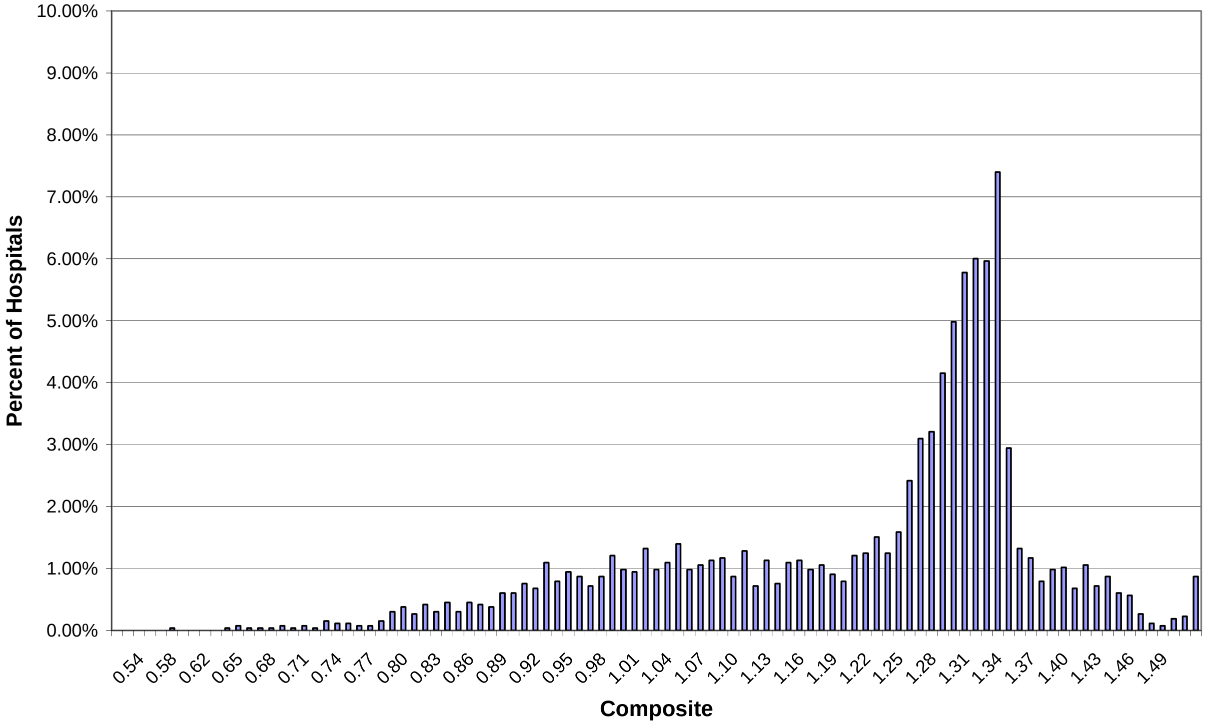


Figure 3.5 - Mortality for Selected Procedures, Factor Weight

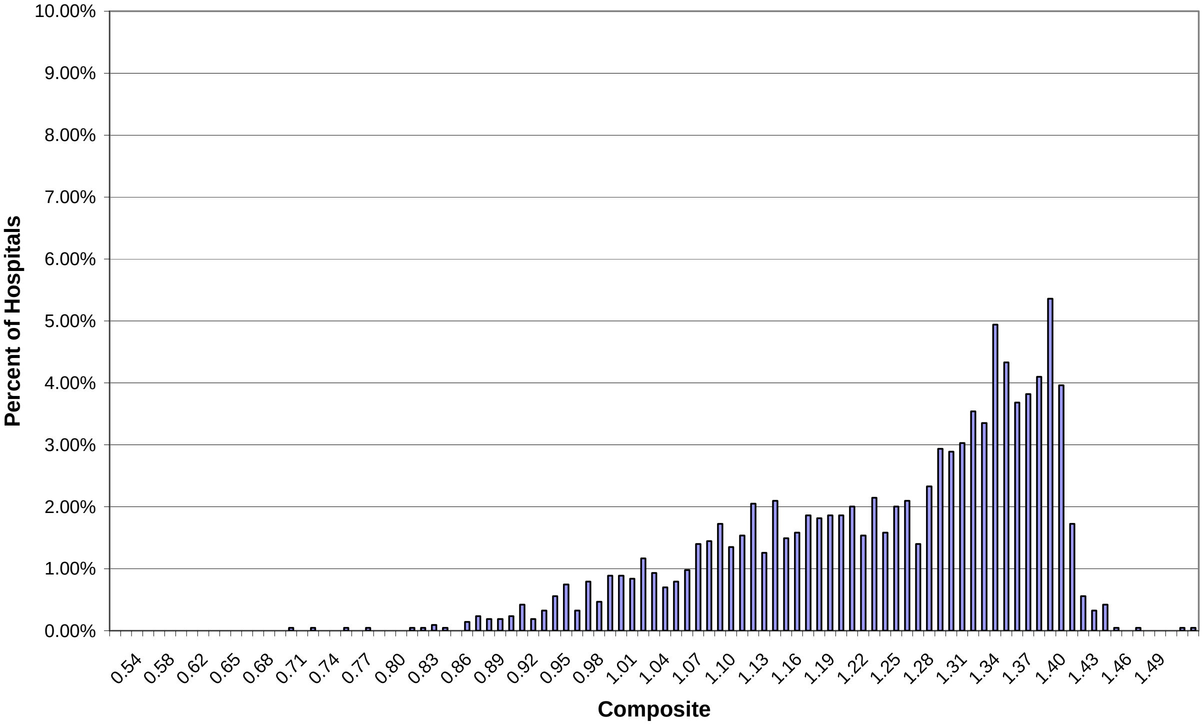


Figure 3.6 - Mortality for Selected Conditions, Single Indicator Weight

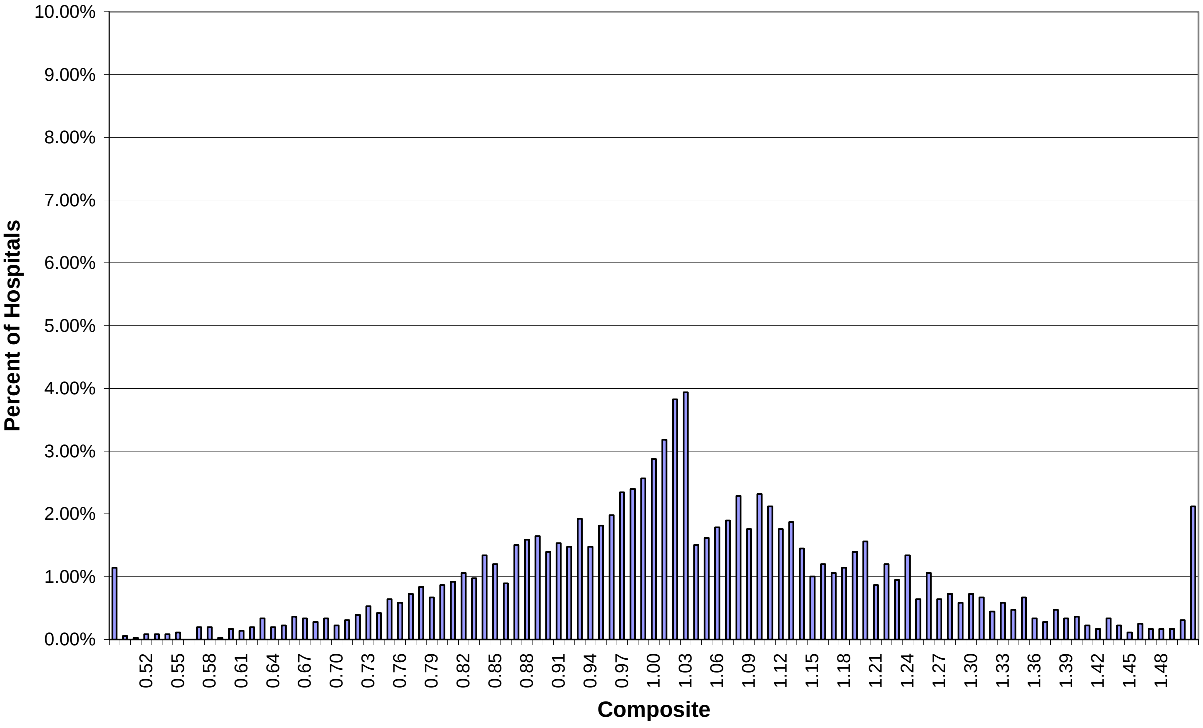


Figure 3.7 - Mortality for Selected Conditions, Equal Weight

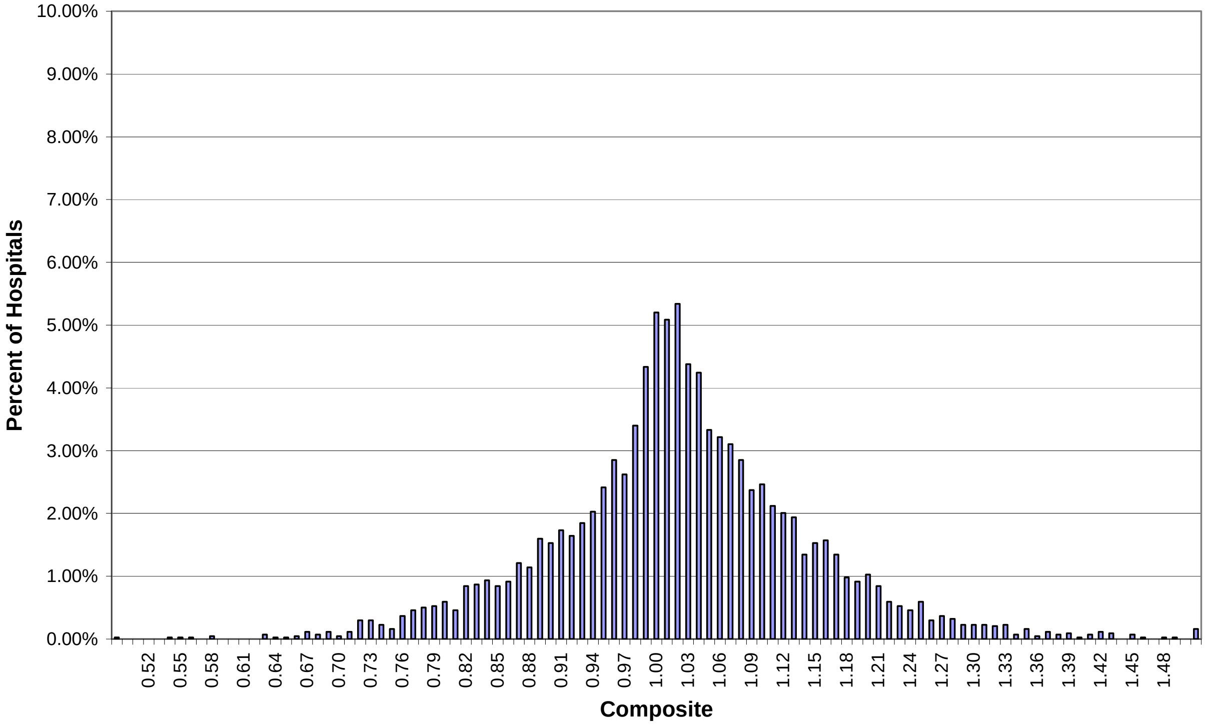


Figure 3.8 - Mortality for Selected Conditions, Numerator Weight

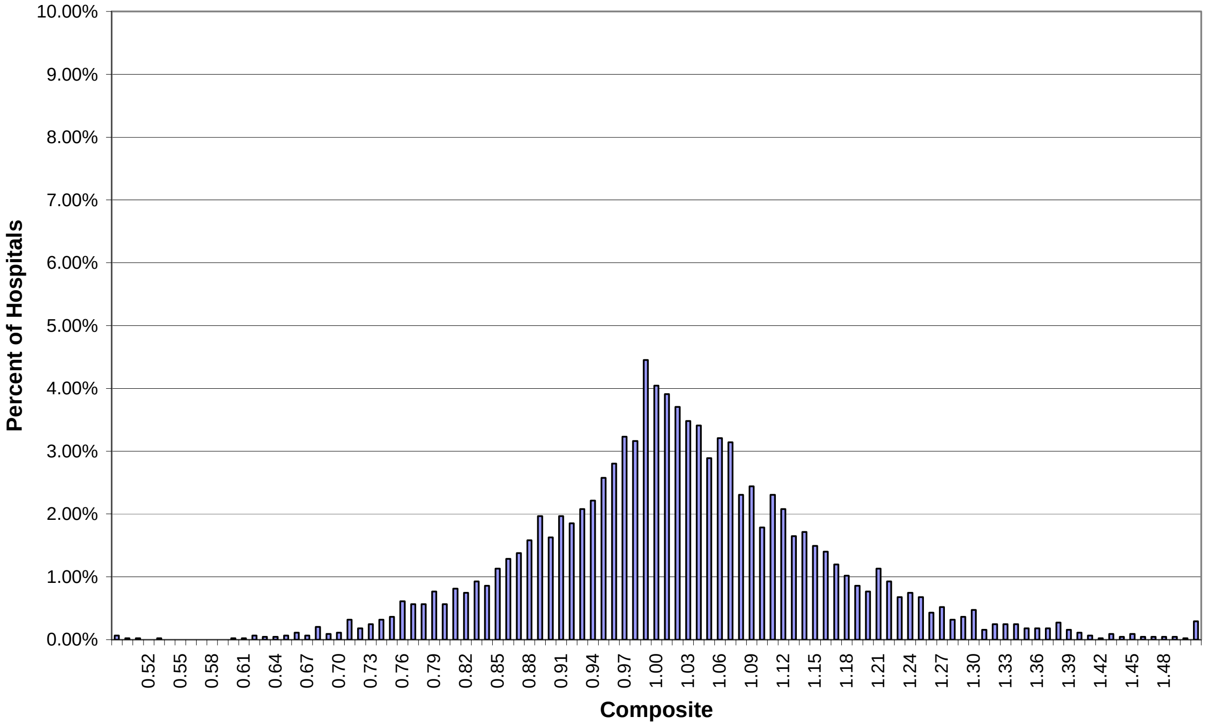


Figure 3.9 - Mortality for Selected Conditions, Denominator Weight

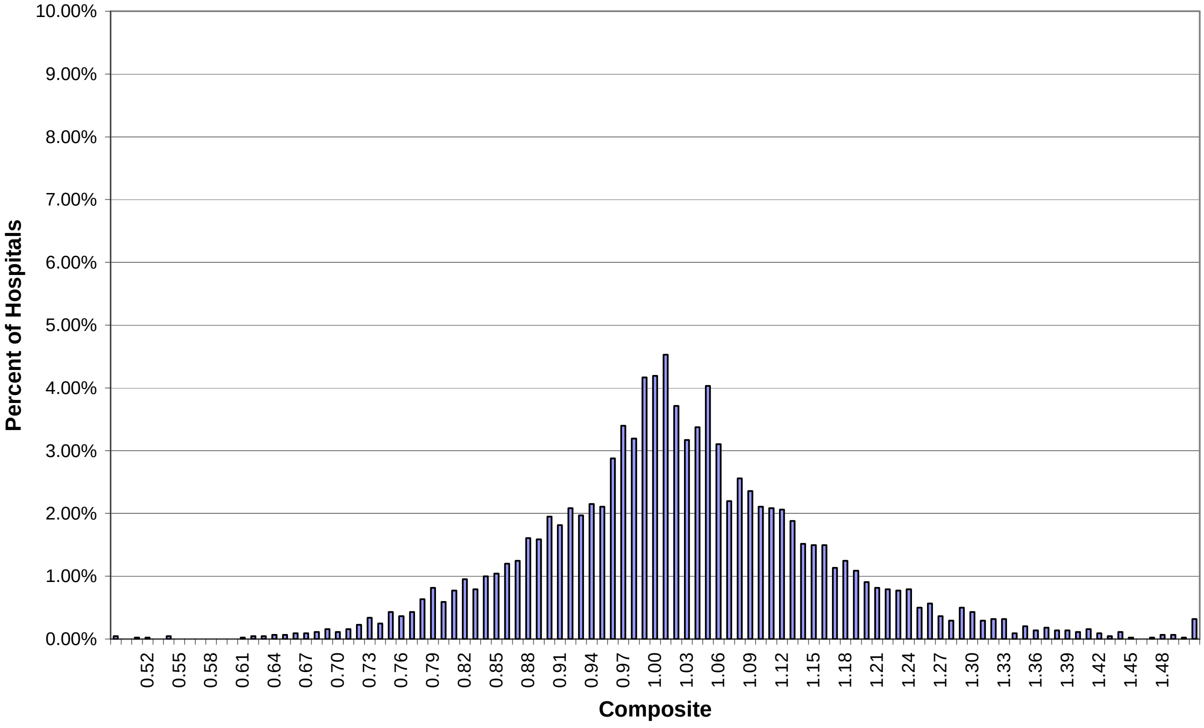
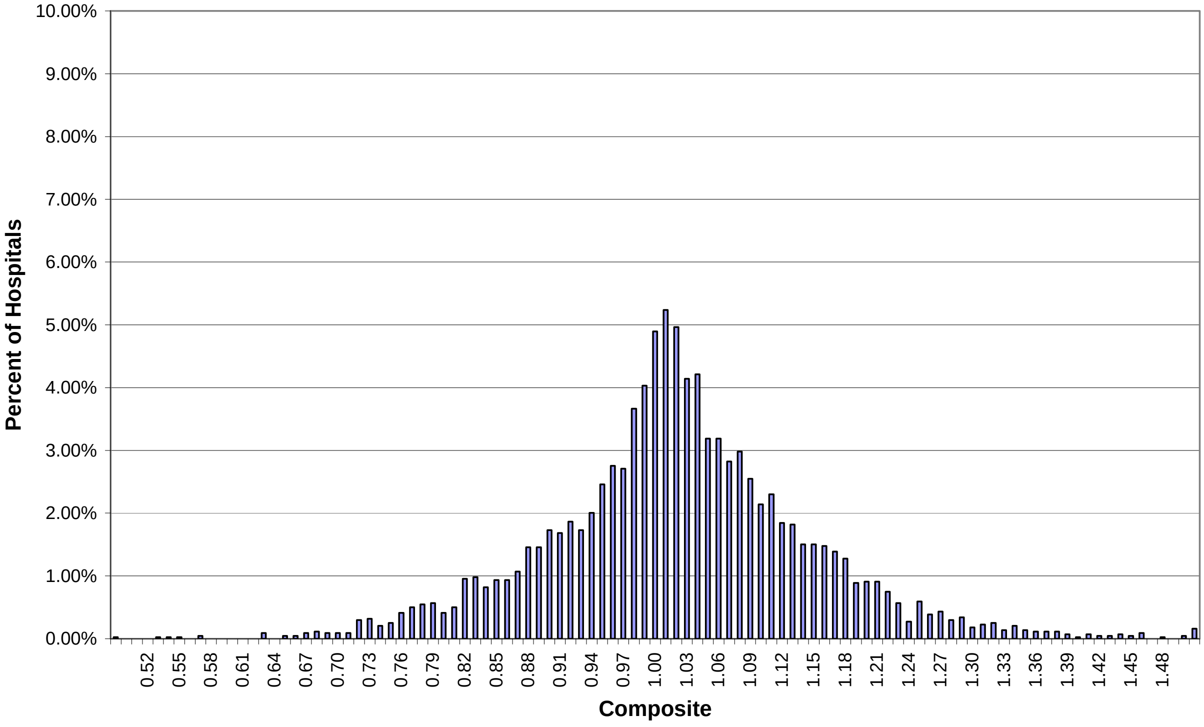


Figure 3.10 - Mortality for Selected Conditions, Factor Weight



Appendix D. Empirical Methods

Introduction

The AHRQ Quality Indicator risk-adjustment modules begin with estimating a simple logistic model of a 0/1 outcome variable and a set of patient-level covariates as dependent variables, and using the results to form the predicted outcome for each patient (e.g., P=pr(outcome=1)).

Notation

Y_{ij} = 0 or 1, outcome for patient j in hospital i
X_{ij} = covariates (e.g., gender, age, DRG, comorbidity)
P_{ij} = predicted probability from logit of Y on X
$= \exp(X_{ij}\beta)/[1+ \exp(X_{ij}\beta)]$
where β is estimated from logit on entire sample
e_{ij} = $Y_{ij} - P_{ij}$ = logit residual (difference between actual and expected)
N = number of patients in sample at hospital i
α = average outcome in the entire sample* (e.g., Y-bar)

* For the AHRQ QI, the sample is the entire reference population consisting of the discharges in the State Inpatient Databases for the participating States pooled over 3 years (2001-2003). Therefore, the "average outcome for the entire sample" is the population rate.

Computing the Noise Variance

Estimate the risk-adjusted ratio (RAR) and noise variance using the Ratio Method (risk-adjusted rate = (observed rate/expected rate) × population rate) of Indirect Standardization for each hospital.

Estimating RAR

Let $O_i = (1/n_i)\sum(Y_{ij})$ be the observed rate at hospital i
Let $E_i = (1/n_i)\sum(P_{ij})$ be the expected rate at hospital i

RAR_i

$= \alpha(O_i/E_i) = \alpha\,[(1/n_i)\sum(Y_{ij})]/[(1/n_i)\sum(P_{ij})]$ (where sum is for $j = 1$ to $j = n_i$)

= population rate × observed/expected at hospital i

Estimating Variance of RAR (standard error is the square root of the variance)

$\mathrm{Var}(RAR_i)$

$= \mathrm{Var}[\alpha(O_i/E_i)]$

$= (\alpha/E_i)^2\mathrm{Var}[O_i]$ (since $\mathrm{var}(aX) = a^2\mathrm{var}(X)$ for any constant a)

$= (\alpha/E_i)^2\mathrm{Var}[(1/n_i)\sum(Y_{ij})]$ (by the definition of O_i)

$= (\alpha/E_i)^2(1/n_i)^2\mathrm{Var}[\sum(Y_{ij})]$ (since $\mathrm{var}(aX) = a^2\mathrm{var}(X)$ for any constant a)

$= (\alpha/E_i)^2(1/n_i)^2[\sum\mathrm{Var}(Y_{ij})]$ (since $\mathrm{var}(\sum X_i)=\sum\mathrm{var}(X_i)$ if X_i are independent)

$= (\alpha/E_i)^2(1/n_i)^2\sum [P_{ij}(1-P_{ij})]$ (since Y is 0/1, $\mathrm{var}(Y) = P(1-P)$)

Computing the Composite Variance

*Setup**

1. Let M be a 1xK vector of observed quality measures (for a given hospital, suppress hospital subscript for convenience), noisy measures of the true underlying 1xK quality vector μ, so that:
    - $M = \mu + \varepsilon$
    - Let the KxK signal variance-covariance be $Var(\mu) = \Omega_\mu$
    - Let the KxK noise variance-covariance be $Var(\varepsilon) = \Omega_\varepsilon$
2. Let $\hat{\mu}$ (1xK) be the posterior (filtered) estimate of μ, so that:
    - $\mu = \hat{\mu} + \nu$, where the 1xK vector ν represents the prediction error of the posterior estimates, and $Var(\nu)$ is the KxK variance-covariance matrix for these posterior estimates.
3. The goal is to estimate the variance for any weighted average of the posterior estimates. For a given (Kx1) weighting vector (w), this is given by:

    $$Var(\nu w) = w' Var(\nu) w$$

    Thus, we simply need an estimate of $Var(\nu)$.

* For more information on the empirical Bayes estimator method, see the technical appendix in Dimick JB, Staiger DO, Birkmeyer JD. Are mortality rates for different operations related?: Implications for measuring the quality of noncardiac surgery. Med Care 2006 Aug;44(8):774-8; and McClellan MB and Staiger DO. The quality of healthcare providers. Cambridge, MA: National Bureau of Economic Research, 1999. NBER Working Paper #7327. Available at: http://www.nber.org/papers/w7327.

Special Case

Filtered estimates are formed in isolation for each measure (univariate) and the estimation error is assumed not correlated across measures (e.g., each measure is based on a different sample of patients or independent patient outcomes).

1. Forming each measure in isolation, using superscripts to indicate the measure (k=1,…,K) as above:

    $$\hat{\mu}^k = M^k \hat{\beta}^k = M^k \left[\Omega_\mu^{kk} + \Omega_\varepsilon^{kk}\right]^{-1} \Omega_\mu^{kk}$$

    $$Var(\nu^k) = \Omega_\mu^{kk} - \Omega_\mu^{kk}\left(\Omega_\mu^{kk} + \Omega_\varepsilon^{kk}\right)^{-1} \Omega_\mu^{kk} = \Omega_\mu^{kk}\left(1 - \hat{\beta}^k\right)$$

    - Note that in this simple case the filtered estimate is a simple shrinkage estimator and:
        - $\hat{\beta}^k$ is the signal ratio of measure k, is the reliability of the measure, and is the r-squared measuring how much of the variation in the true measure can be explained with the filtered measure.
        - The variance of the filtered estimate is simply the signal variance times 1 minus the signal ratio. Thus, if the signal ratio is 0 (no information in the measure), the error in the estimate is equal to the signal variance. But as the signal ratio grows, the error in the estimate shrinks (to 0 if there is a signal ratio of 1 – no noise).

2. The formula for $Var(\nu^k)$ above provides the diagonal elements of $Var(\nu)$ (the full KxK variance-covariance matrix of the filtered estimates). So, one gets the covariance elements, which are (for j≠k):

   $$Cov(\nu^j, \nu^k) = E[(\mu^j - \hat{\mu}^j)(\mu^k - \hat{\mu}^k)]$$

   - After some algebra (assuming independent estimation error in the two measures), one gets the following simple expression:

     $$Cov(\nu^j, \nu^k) = \Omega_\mu^{jk}(1 - \hat{\beta}^j)(1 - \hat{\beta}^k)$$

   - Note that this is just the signal covariance times 1 minus the signal ratio for each of the measures. Thus, if the signal ratio is 0 for each measure, the covariance in the estimates is simply the signal covariance. As either measure gets a stronger signal ratio (becomes more precise), the covariance in the estimates shrinks to 0.
   - Also note that if one measure is missing, then the signal ratio is simply set to 0. The filtered estimate is shrunk all the way back to the (conditional) mean, and the variance and covariance are as defined above.