AHRQ Quality Indicators
Inpatient Quality Indicators (IQI)
Composite Measure

1. Introduction

Many users of the AHRQ Quality Indicators (AHRQ QI) 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] 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 construct 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: *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. In order 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.

The content of this report is very technical in nature. In order to facilitate future use of the composite, the AHRQ QI program will be developing 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 criteria for selection, coding and specifications, see the Guide to Inpatient Quality Indicators and the IQI Technical Specifications, available on the AHRQ QI website[3].

2. Why 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

---

[1] The most recent National Healthcare Quality Report and National Healthcare Disparities Report may be found at http://qualitytools.ahrq.gov.

[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_v30.pdf; Technical Specifications: http://www.qualityindicators.ahrq.gov/downloads/iqi/iqi_technical_specs_v30.pdf.

Composite measures have several potential benefits over individual indicators.

Summarize quality across multiple indicators. There are thirty (30) provider-level Inpatient Quality Indicators for various conditions and procedures, making it difficult to formulate general statements about overall trends or differences in quality.

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 structure 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) healthcare needs. Depending on how the component indicators are weighted, composites may reflect the likely health outcomes for an individual or population.

Avoid cognitive "short-cuts". Research suggests that individuals faced with too many factors in making a decision take cognitive short-cuts 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.

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.

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

Identify which parts of the healthcare 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.

Detract from the impact and credibility of reports. The composite measure might not reflect the evidence-base of the individual indicators.

2.3 Potential Uses of Composite Measures.

Composite measures have many potential uses.

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

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

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

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

3. Alternative Perspectives on Composite Measures

There are two alternative perspectives on composite measures that guide the development of a composite measure methodology.

Signaling perspective. The signaling perspective seeks to guide decision making 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 decision making on allocating investment resources.

Psychometric perspective. The psychometric perspective 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 individual 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 decision making 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 Inpatient Quality Indicators: *Mortality for Selected Procedures* and *Mortality for Selected Conditions*. 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 Inpatient Quality Indicators, 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, 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 each other, 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 #15 In-Hosp Mort AMI
IQI #09 In-Hosp Mort Pancreatic Resection	IQI #16 In-Hosp Mort CHF
IQI #11 In-Hosp Mort AAA Repair	IQI #17 In-Hosp Mort Stroke
IQI #12 In-Hosp Mort CABG	IQI #18 In-Hosp Mort GI Hemorrhage
IQI #13 In-Hosp Mort Craniotomy	IQI #19 In-Hosp Mort Hip Fracture
IQI #14 In-Hosp Mort Hip Replacement	IQI #20 In-Hosp Mort Pneumonia
IQI #30 In-Hosp Mort PTCA	
IQI #31 In-Hosp Mort Carotid Endarterectomy	

4.2 The 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 individual indicators, where the weights are selected based on the intended use of the measure. These weights might be determined empirically or based on non-empirical considerations.

4.3 Constructing the AHRQ QI Composite Measure

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, and then summing the predicted value among all the cases in the hospital to compute the expected rate. The risk-adjusted rate is computed using indirect standardization as the observed rate divided by the expected rate multiplied by the reference population rate. The current reference population is the states participating in the

---

[4] A separate workgroup is evaluating alternative risk-adjustment and hierarchical modeling methodologies for the AHRQ QI.

HCUP program for 2001-2003, consisting of approximately 38 states and 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 of the Inpatient Quality Indicators. The levels of the rates vary from indicator to indicator. In order to combine the individual indicators using a common scale, each indicator is first divided by the reference population rate. The components of the composite are therefore defined as deviations (i.e. a ratio) from the overall mean for each indicator.

Step3. Compute the reliability-adjusted ratio.

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

reliability-adjusted ratio = [risk-adjusted ratio * weight] + [reference population ratio * (1 – weight)]

Table 5 shows the average reliability weights for the Inpatient Quality Indicators based on denominator size. For small providers, the weight is closer to zero. 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

---

[5] The state data organizations that participated in the 2001-03 HCUP SID: Arizona Department of Health Services; California Office of Statewide Health Planning & Development; Colorado Health & Hospital Association; Connecticut - Chime, Inc.; Florida Agency for Health Care Administration; Georgia: An Association of Hospitals & 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; 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 & Senior Services; New York State Department of Health; North Carolina Department of Health and Human Services; Ohio Hospital Association; Oregon Association of Hospitals & Health Systems; Pennsylvania Health Care Cost Containment Council; Rhode Island Department of Health; South Carolina State 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 & Family Services.

The composite measure is the weighted average of the scaled and reliability-adjusted ratios for the individual 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.

Equal weight. In this case, each component indicator is assigned an identical weight based on the number of indicators. That is, the weight is equal to one divided by the number of indicators in the composite.

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 for experiencing the outcome of interest for 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 individual patient.

Factor weight. A factor weight is based on some sort of analysis which 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 * weight1] + [indicator2 * weight2] + . . . + [indicatorN * weightN]

The confidence interval on 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 the appendix.

4.4 An example 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 Inpatient Quality Indicator. 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 should be able to trace the computation from the individual indicators to the composite and back again.

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

Table S1. A Single Provider in a Single Year

IQI	Patients	Observed Rate	Risk-adjusted Rate	Risk-adjusted SE
IQI #15 In-Hosp Mort AMI	55	223.4	145.3	30.0
IQI #16 In-Hosp Mort CHF	419	88.6	119.0	12.0
IQI #17 In-Hosp Mort Stroke	133	160.7	285.4	31.9
IQI #18 In-Hosp Mort GI Hemorrhage	193	51.3	52.0	12.5
IQI #19 In-Hosp Mort Hip Fracture	90	59.7	83.3	21.6
IQI #20 In-Hosp Mort Pneumonia	664	115.7	146.1	13.5

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

Table S2. Scaling the Single Provider Rate

IQI	Reference Population Rate	Risk-adjusted Ratio	Risk-adjusted SE
IQI #15 In-Hosp Mort AMI	93.5	1.5539	0.3208
IQI #16 In-Hosp Mort CHF	45.69	2.6056	0.2631
IQI #17 In-Hosp Mort Stroke	113.29	2.5191	0.2812
IQI #18 In-Hosp Mort GI Hemorrhage	31.59	1.6468	0.3958
IQI #19 In-Hosp Mort Hip Fracture	32.15	2.5896	0.6719
IQI #20 In-Hosp Mort Pneumonia	83.09	1.7582	0.1622

The individual 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

Table S3A. Compute the Reliability Weight

IQI	Risk-adjusted	Noise Variance	Signal Variance	Reliability Weight

	SE			
IQI #15 In-Hosp Mort AMI	0.3208	0.1029	0.0422	0.2909
IQI #16 In-Hosp Mort CHF	0.2631	0.0692	0.1418	0.6719
IQI #17 In-Hosp Mort Stroke	0.2812	0.0791	0.0922	0.5384
IQI #18 In-Hosp Mort GI Hemorrhage	0.3958	0.1567	0.0751	0.3241
IQI #19 In-Hosp Mort Hip Fracture	0.6719	0.4515	0.1611	0.2630
IQI #20 In-Hosp Mort Pneumonia	0.1622	0.0263	0.0721	0.7326

Note: Noise variance is standard error squared (for details on calculating the noise variance, see Appendix D.2); 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.

Table S3B. Compute the Reliability-adjusted Ratio

IQI	Reliability Weight	Risk-adjusted Ratio	Reference Population Ratio	Reliability-adjusted Ratio
IQI #15 In-Hosp Mort AMI	0.2909	1.5539	1.0000	1.1611
IQI #16 In-Hosp Mort CHF	0.6719	2.6056	1.0000	2.0788
IQI #17 In-Hosp Mort Stroke	0.5384	2.5191	1.0000	1.8179
IQI #18 In-Hosp Mort GI Hemorrhage	0.3241	1.6468	1.0000	1.2096
IQI #19 In-Hosp Mort Hip Fracture	0.2630	2.5896	1.0000	1.4181
IQI #20 In-Hosp Mort Pneumonia	0.7326	1.7582	1.0000	1.5555

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.

Step 4. Select the component weights

Table S4. Numerator Weight

	Denominator Weight
IQI #15 In-Hosp Mort AMI	0.1434
IQI #16 In-Hosp Mort CHF	0.2554
IQI #17 In-Hosp Mort Stroke	0.1246
IQI #18 In-Hosp Mort GI Hemorrhage	0.1152
IQI #19 In-Hosp Mort Hip Fracture	0.0691
IQI #20 In-Hosp Mort Pneumonia	0.2924

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

Step 5. Construct the composite measure

Table S5. Construct the Composite Measure

	Denominator Weight (A)	**Reliability-adjusted Ratio (B)**	**(A) *(B)**
IQI #15 In-Hosp Mort AMI	0.1434	1.1611	0.1665
IQI #16 In-Hosp Mort CHF	0.2554	2.0788	0.5309
IQI #17 In-Hosp Mort Stroke	0.1246	1.8179	0.2265
IQI #18 In-Hosp Mort GI Hemorrhage	0.1152	1.2096	0.1393
IQI #19 In-Hosp Mort Hip Fracture	0.0691	1.4181	0.0980
IQI #20 In-Hosp Mort Pneumonia	0.2924	1.5555	0.4548
Mortality for Selected Conditions			*1.6161*
Standard Error			*0.07195*
Confidence Interval at p<0.05		*1.7571*	*1.4751*

Note: For details on calculating the composite variance (standard error), see Appendix D.3

The final composite is simply the weighted average of the component indicators. Note the potential application of the composite construction for use in quality improvement. The final computation shows that CHF 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 CHF mortality as the best mechanism to lowering the composite score.

5. Performance of the AHRQ QI Composite Measures

5.1 Evaluation Criteria

The tables and figures 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 composite over time. 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 percent of providers that are either worse than average or better than average based on the confidence interval for the composite measure. The discrimination performance varies depending on the weight used. The single indicator weights have 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, the factor weight and equal weight.

The *Mortality for Selected Procedures* composite measure has more variability in the ability to discriminate performance, 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. Figure 4 shows the distribution of each composite and the 95 percent confidence interval.

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 2001-2002. The columns show the relative difference in the actual risk-adjusted ratio in 2003 for the best and worst performing quintile relative to the middle sixty 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.

Table 9 shows the correlation among the composite measures using the alternative weights. For the *Mortality for Selected Procedures*, the correlations vary from 0.70 to 0.80. For the *Mortality for Selected Conditions*, the correlations are very high regardless of the weight.

The table also shows the correlation in the composite measures from one year to the next. For the *Mortality for Selected Procedures*, the correlation depends on the weight used, with the numerator weight showing the most persistence. For the *Mortality for Selected Conditions*, the correlation does not depend much on the weight used.

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 reflect in the individual indicators themselves, and that the composites are only as useful and valid as are the individual indicators that make up the composite. The AHRQ QI are currently undergoing review through the NQF consensus development processes, and that as part of that process 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 Quality Indicators and the data upon which they are based continue to improve, the composite measures will improve as potentially useful tools for decision making 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, 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
- Bob Houchens, Medstat
- Bill Rogers, Rogers Associate
- Chunliu Zhan, Agency for Healthcare Research and Quality

AHRQI 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 Data, 2001-2003; Rate per 1,000

Table 2. Provider-level Rates

		Risk adjusted		Reliability adjusted	
IQI	**Providers**	**Rate**	**Std. Dev.**	**Rate**	**Std. Dev.**
IQI #08 In-Hosp Mort Esophageal Resection	424	99.44	198.00	79.88	34.76
IQI #09 In-Hosp Mort Pancreatic Resection	857	77.56	168.36	63.18	29.24
IQI #11 In-Hosp Mort AAA Repair	1,963	91.01	103.63	79.50	16.13
IQI #12 In-Hosp Mort CABG	1,021	36.77	18.46	35.85	9.18
IQI #13 In-Hosp Mort Craniotomy	1,551	73.33	69.77	71.99	11.09
IQI #14 In-Hosp Mort Hip Replacement	3,064	3.47	19.73	2.88	3.92
IQI #30 In-Hosp Mort PTCA	1,261	14.15	10.87	13.86	2.32
IQI #31 In-Hosp Mort Carotid Endarterectomy	2,339	9.61	66.12	7.01	4.57
IQI #15 In-Hosp Mort AMI	3,962	108.88	65.77	96.15	13.79
IQI #16 In-Hosp Mort CHF	4,369	55.24	58.29	48.42	15.05
IQI #17 In-Hosp Mort Stroke	4,230	130.00	100.19	116.94	27.04
IQI #18 In-Hosp Mort GI Hemorrhage	4,162	33.44	41.61	31.94	4.71
IQI #19 In-Hosp Mort Hip Fracture	3,583	35.47	46.16	33.36	8.26
IQI #20 In-Hosp Mort Pneumonia	4,463	85.57	52.98	84.74	19.57

Source: HCUP State Inpatient Data, 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 Data, 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 Data, 2001-2003

Table 5. Reliability Weight by Denominator Size

IQI	Providers	Q1	Q2	Q3	Q4
Average Annual Denominator Size (by quartile)					
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

IQI	Q1	Q2	Q3	Q4	Weighted Average
Average Reliability Weight					
IQI #08 In-Hosp Mort Esophageal Resection	0.2376	0.3098	0.3896	0.6274	0.5710
IQI #09 In-Hosp Mort Pancreatic Resection	0.2437	0.3187	0.4137	0.6435	0.6135
IQI #11 In-Hosp Mort AAA Repair	0.1549	0.2924	0.4441	0.6336	0.5917
IQI #12 In-Hosp Mort CABG	0.3492	0.6337	0.7551	0.8689	0.8016
IQI #13 In-Hosp Mort Craniotomy	0.0728	0.2489	0.4348	0.6679	0.6562
IQI #14 In-Hosp Mort Hip Replacement	0.1481	0.3560	0.5714	0.7738	0.7088
IQI #30 In-Hosp Mort PTCA	0.1118	0.3915	0.5783	0.7395	0.6779
IQI #31 In-Hosp Mort Carotid Endarterectomy	0.1672	0.4388	0.6543	0.8371	0.7593
IQI #15 In-Hosp Mort AMI	0.1018	0.3515	0.6331	0.8276	0.8021
IQI #16 In-Hosp Mort CHF	0.2199	0.5587	0.8012	0.9173	0.8591
IQI #17 In-Hosp Mort Stroke	0.1861	0.5082	0.7752	0.9245	0.8577
IQI #18 In-Hosp Mort GI Hemorrhage	0.0556	0.2194	0.4438	0.6652	0.5718
IQI #19 In-Hosp Mort Hip Fracture	0.1335	0.3739	0.5713	0.7366	0.6447
IQI #20 In-Hosp Mort Pneumonia	0.2661	0.6279	0.8336	0.9248	0.8601

Source: HCUP State Inpatient Data, 2001-2003

Table 6. Alternative Composite Weights

	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	**0.1388**
IQI #09 In-Hosp Mort Pancreatic Resection	**1.0000**	0.1250	0.0087	0.0031	0.1315
IQI #11 In-Hosp Mort AAA Repair	0.0000	0.1250	0.0922	0.0260	**0.1933**
IQI #12 In-Hosp Mort CABG	0.0000	0.1250	**0.3355**	0.2207	0.1917
IQI #13 In-Hosp Mort Craniotomy	0.0000	0.1250	**0.2359**	0.0728	0.0166
IQI #14 In-Hosp Mort Hip Replacement	0.0000	**0.1250**	0.0151	0.1202	0.0101
IQI #30 In-Hosp Mort PTCA	0.0000	0.1250	0.2808	**0.4645**	0.2010
IQI #31 In-Hosp Mort Carotid Endarterectmy	0.0000	**0.1250**	0.0280	0.0916	0.1169
IQI #15 In-Hosp Mort AMI	0.0000	0.1667	**0.1933**	0.1434	0.1595
IQI #16 In-Hosp Mort CHF	0.0000	0.1667	0.1683	**0.2554**	0.1790
IQI #17 In-Hosp Mort Stroke	0.0000	0.1667	**0.2036**	0.1246	0.1777
IQI #18 In-Hosp Mort GI Hemorrhage	0.0000	**0.1667**	0.0525	0.1152	0.1464
IQI #19 In-Hosp Mort Hip Fracture	**1.0000**	0.1667	0.0320	0.0691	0.1437
IQI #20 In-Hosp Mort Pneumonia	0.0000	0.1667	**0.3504**	0.2924	0.1937

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

Table 7. Discrimination Performance of Alternative Composites

Composite	Providers	%Better than Average	%Worse than Average	% Total
Mortality for Selected Procedures				
Single Indicator Weight	857	0.58%	2.33%	2.91%
Equal Weight	2,688	0.82%	2.57%	3.39%
Numerator Weight	1,978	4.75%	9.35%	14.10%
Denominator Weight	2,649	1.62%	8.27%	9.89%
Factor Weight	2,116	0.76%	3.12%	3.88%
Mortality for Selected Conditions				
Single Indicator Weight	3,583	3.13%	3.60%	6.73%
Equal Weight	4,384	10.70%	11.36%	22.06%
Numerator Weight	4,426	13.78%	15.95%	29.73%
Denominator Weight	4,415	13.27%	14.88%	28.15%
Factor Weight	4,393	11.47%	12.41%	23.88%

Source: HCUP State Inpatient Data, 2001-2003

Table 8. Forecast 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%	-0.137	**-0.529***	0.160	-0.044	-0.336*	-0.467	-0.180*	-0.075
Worst 20%	0.492**	**0.492***	0.149	-0.067	0.088**	-0.540	0.043	0.217
Equal Weight								
Best 20%	-0.441	-0.035	-0.146	-0.121*	-0.123*	**-0.567****	-0.149*	**0.016**
Worst 20%	-0.197	0.079	0.454*	0.145*	0.123*	**-0.039**	0.050	**0.362***
Numerator Weight								
Best 20%	-0.300	-0.273	-0.175	**-0.171***	**-0.162***	-0.522	-0.175*	-0.019
Worst 20%	0.180	-0.389	0.515*	**0.394***	**0.208***	0.012	0.234*	0.364*
Denominator Weight								
Best 20%	-0.163	-0.082	-0.101	-0.193*	-0.091*	0.302	**-0.145***	-0.077
Worst 20%	1.013*	0.068	0.603*	0.281*	0.196*	0.050	**0.234***	0.390*
Factor Weight								
Best 20%	**-0.354**	0.116	**-0.104**	-0.107*	-0.195*	-0.443	-0.161*	0.007
Worst 20%	**-0.147**	-0.257	**0.544***	0.275*	0.148*	0.005	0.194*	0.444*

Source: HCUP State Inpatient Data, 2001-2003; *Significant at p<.05; ** Significant at p<.10. The forecast predicts performance in 2003 (ratio) based on performance in 2001-2002 (by quintile) using five alternative measures composite weights. For each indicator, the most highly weighted composite is in **bold**.

Table 8 (continued). Forecast 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%	-0.185*	-0.195*	-0.268*	-0.127*	**-0.312***	-0.182*
Worst 20%	0.083*	0.167*	0.116*	0.163*	**0.249***	0.141*
Equal Weight						
Best 20%	-0.231*	-0.273*	-0.252*	**-0.216***	-0.339*	-0.231*
Worst 20%	0.183*	0.316*	0.189*	**0.242***	0.258*	0.263*
Numerator Weight						
Best 20%	**-0.206***	-0.280*	**-0.252***	-0.195*	-0.330*	**-0.241***
Worst 20%	**0.222***	0.333*	**0.213***	0.264*	0.285*	**0.290***
Denominator Weight						
Best 20%	-0.213*	**-0.284***	-0.239*	-0.212*	-0.318*	-0.238*
Worst 20%	0.203*	**0.331***	0.199*	0.248*	0.311*	0.280*
Factor Weight						
Best 20%	-0.230*	-0.276*	-0.255*	-0.209*	-0.342*	-0.235*
Worst 20%	0.186*	0.318*	0.190*	0.244*	0.260*	0.263*

Source: HCUP State Inpatient Data, 2001-2003; *Significant at p<.05; ** Significant at p<.10. The forecast predicts performance in 2003 (ratio) based on performance in 2001-2002 (by quintile) using five alternative measures composite weights. For each indicator, the most highly weighted composite is in **bold**.

Table 9. Correlation of Alternative Composites

Composite	Single Indicator Weight	Equal Weight	Numerator Weight	Denominator Weight	Factor Weight	Year-to-Year
Mortality for Selected Procedures						
Single Indicator Weight	1.000	0.492*	0.193*	0.159*	0.554*	0.282
Equal Weight		1.000	0.668*	0.846*	0.843*	0.200
Numerator Weight			1.000	0.837*	0.786*	0.792
Denominator Weight				1.000	0.712*	0.341
Factor Weight					1.000	0.613
Mortality for Selected Conditions						
Single Indicator Weight	1.000	0.704*	0.559*	0.590*	0.680*	0.306
Equal Weight		1.000	0.971*	0.981*	0.999*	0.717
Numerator Weight			1.000	0.992*	0.980*	0.733
Denominator Weight				1.000	0.988*	0.726
Factor Weight					1.000	0.726

Source: HCUP State Inpatient Data, 2001-2003; Correlation in alternative composites between 2002 and 2003.

Appendix C. IQI Composite Figures

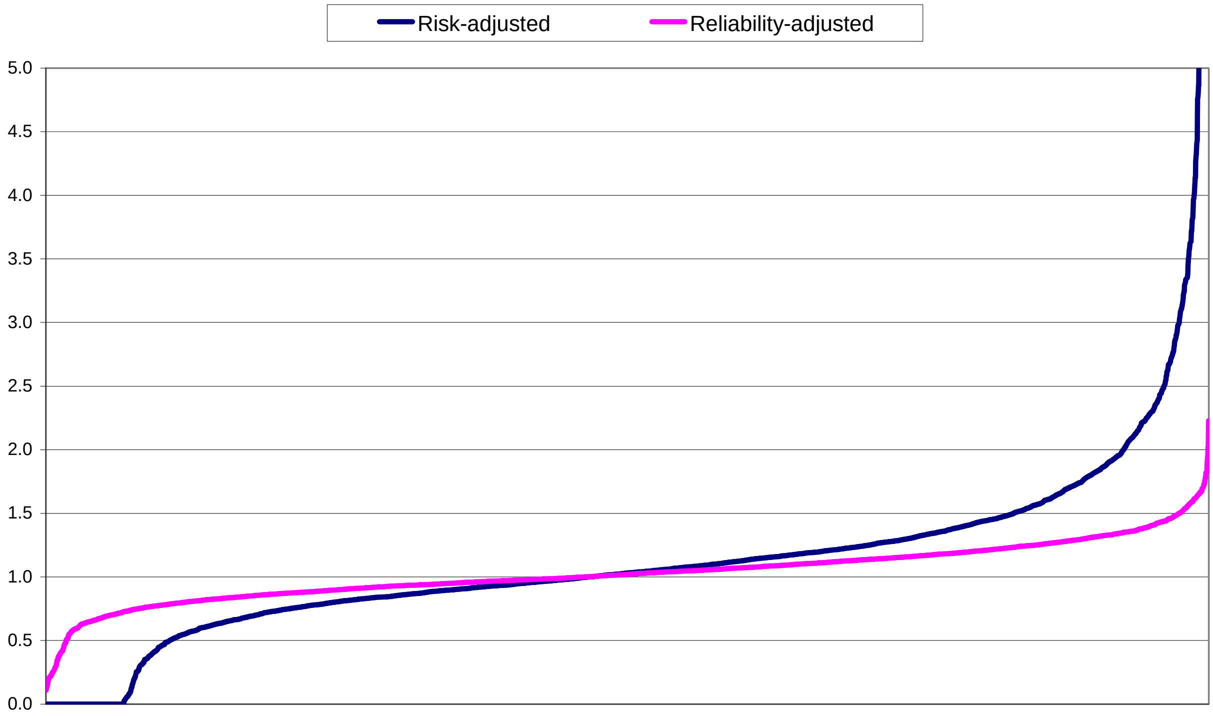


Figure 2. Impact of Reliability Weight for a Single Provider
In-Hosp Mort Stroke

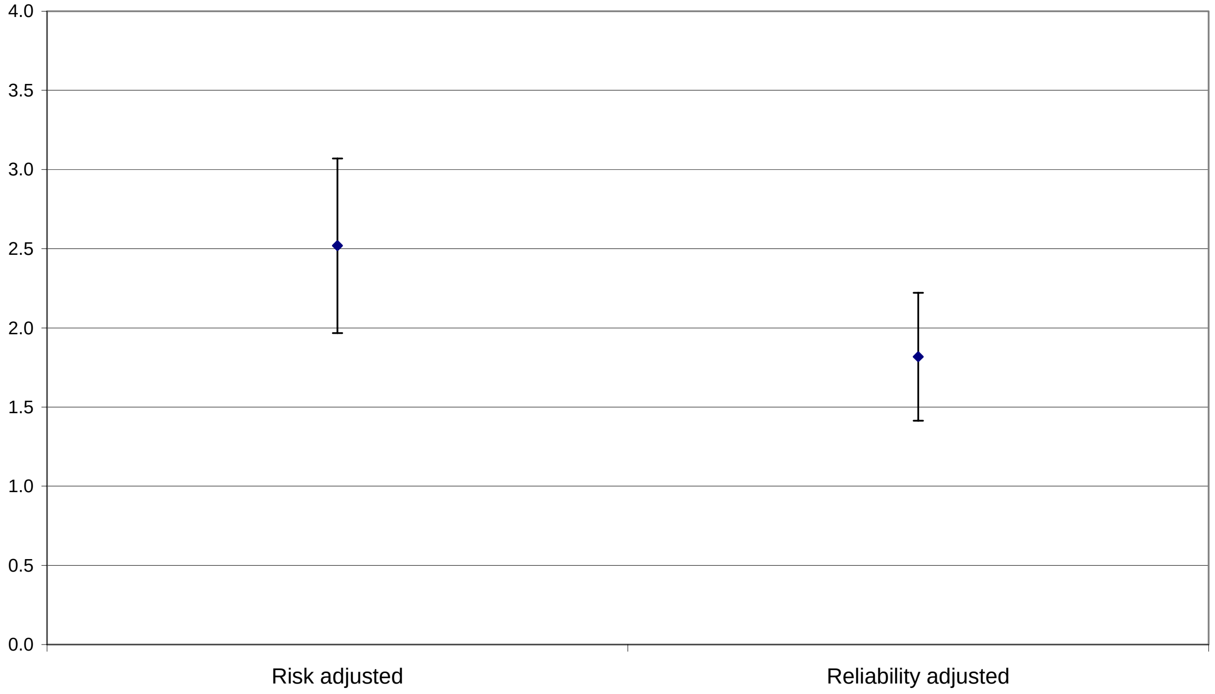


Figure 3. Impact of Reliablity Weight by Denominator Size
In-Hosp Mort Stroke

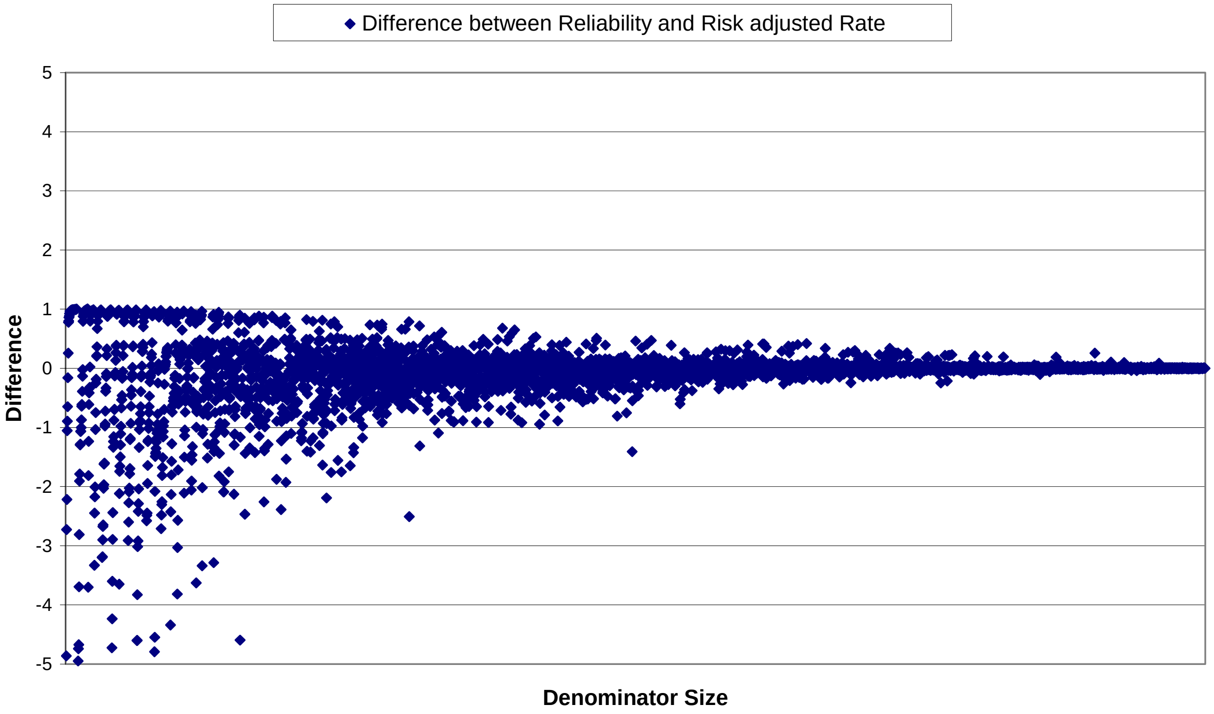


Figure 4-1. Mortality for Selected Procedures, Single Indicator Weight

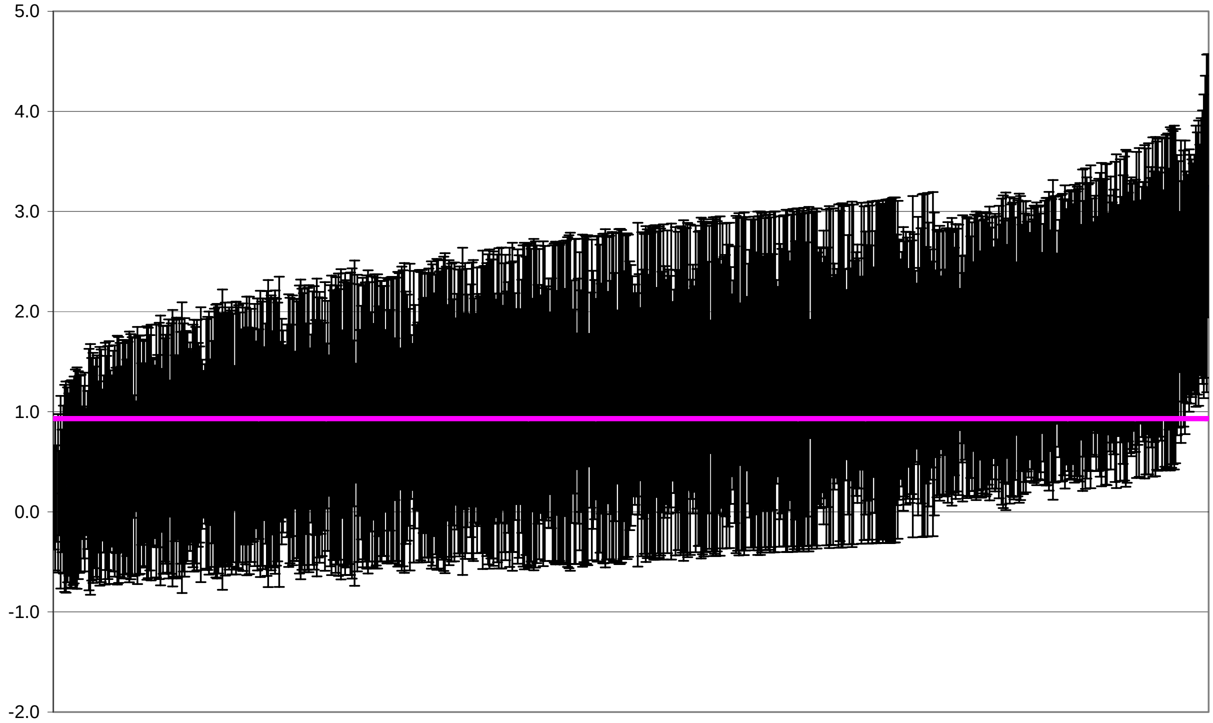


Figure 4-2. Mortality for Selected Procedures, Equal Weight

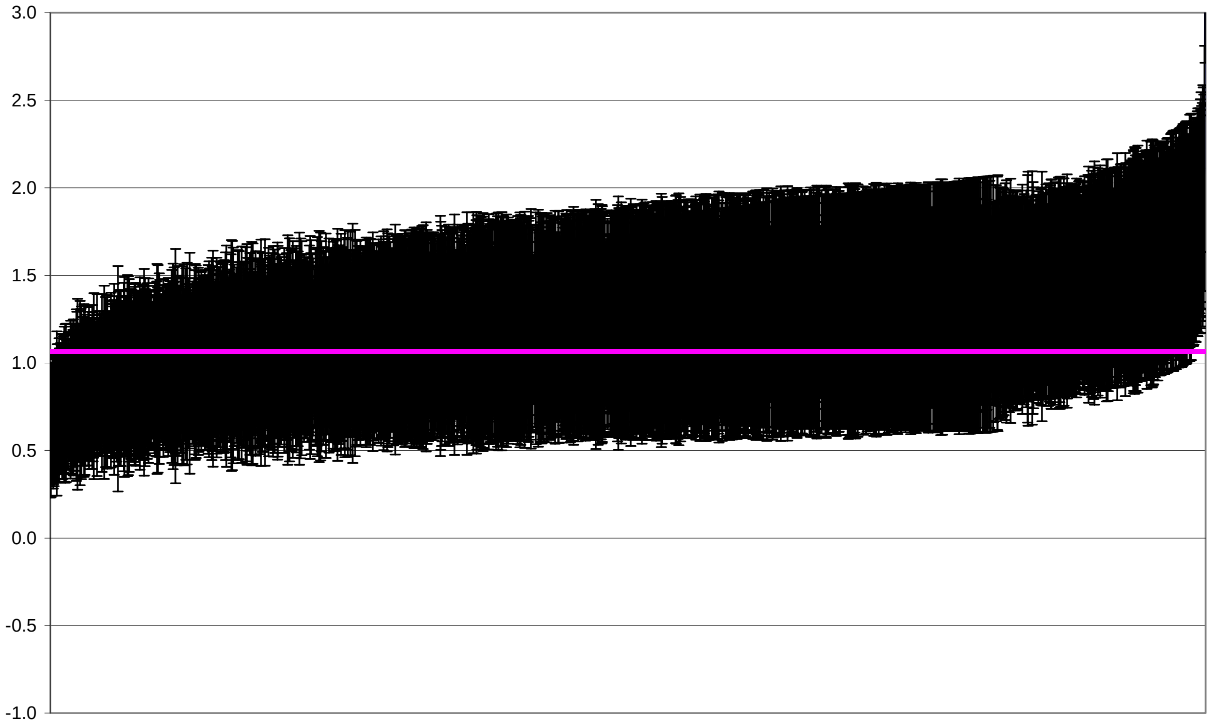


Figure 4-3. Mortality for Selected Procedures, Numerator Weight

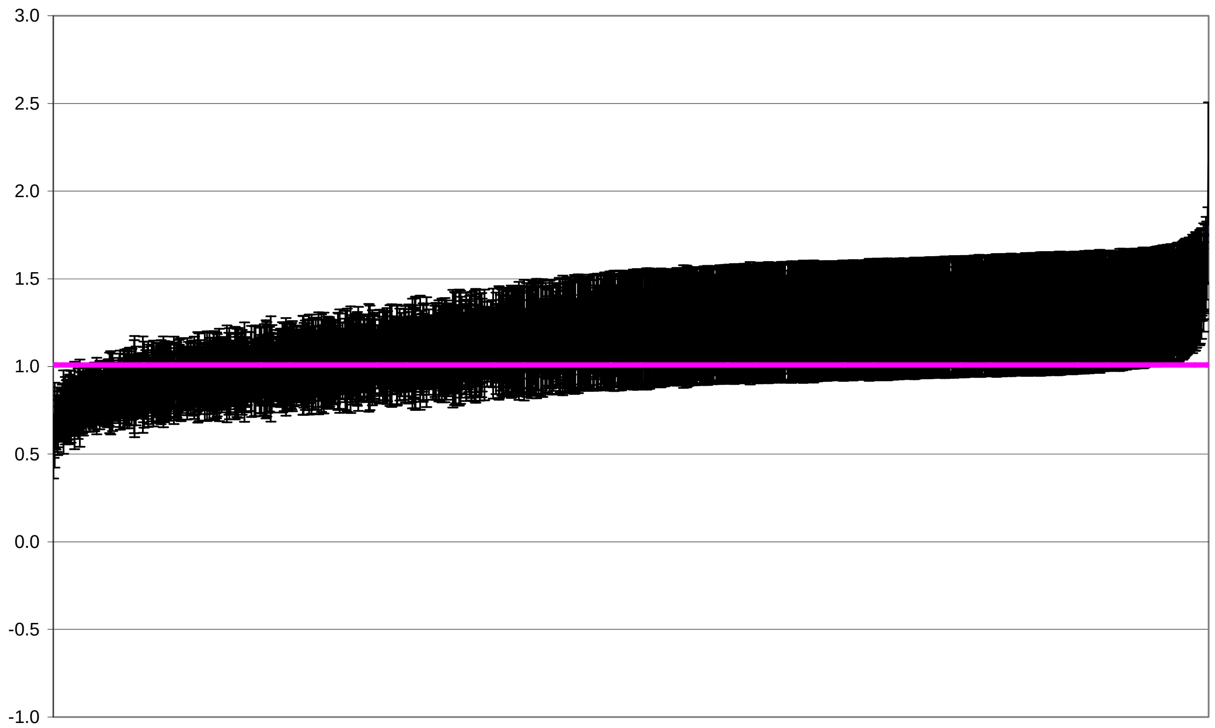


Figure 4-4. Mortality for Selected Procedures, Denominator Weight

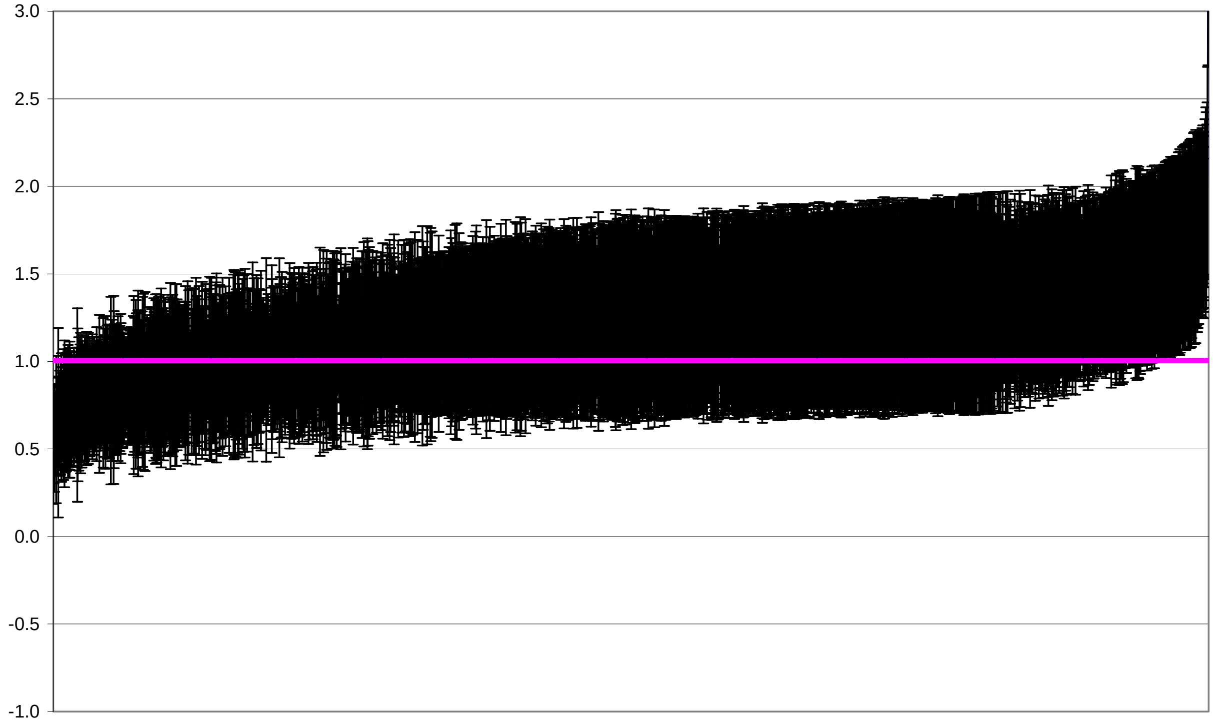


Figure 4-5. Mortality for Selected Procedures, Factor Weight

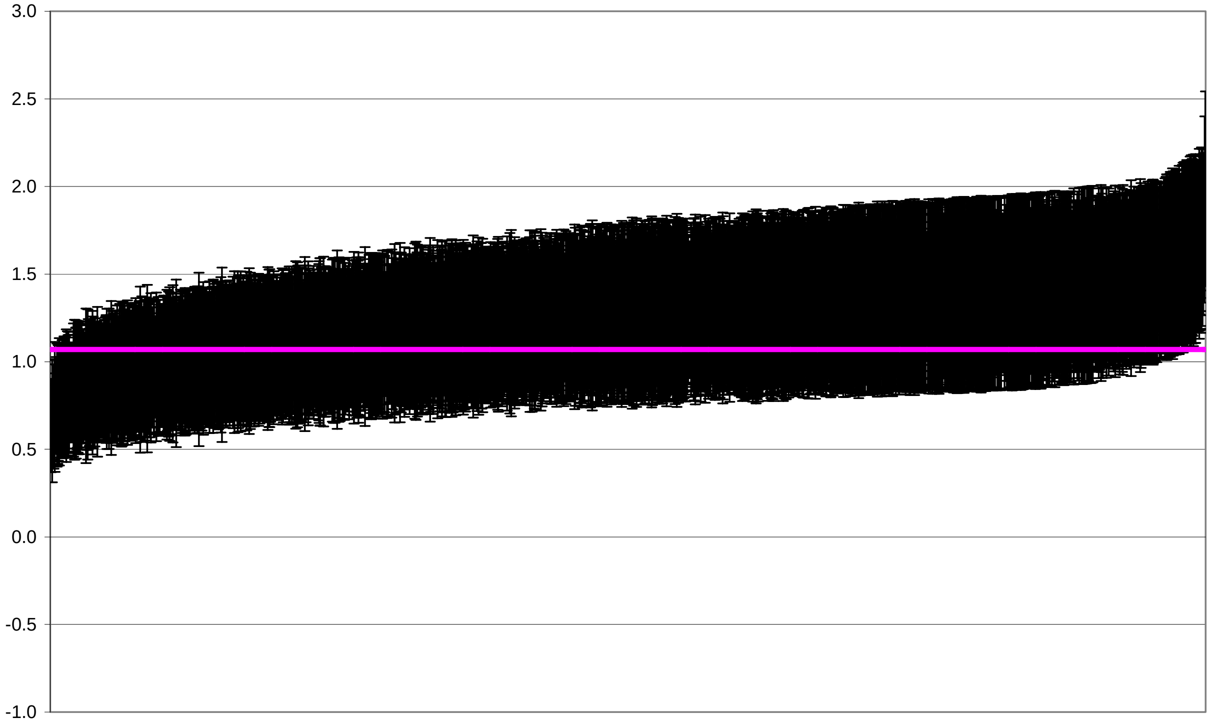


Figure 4-6. Mortality for Selected Conditions, Single Indicator Weight

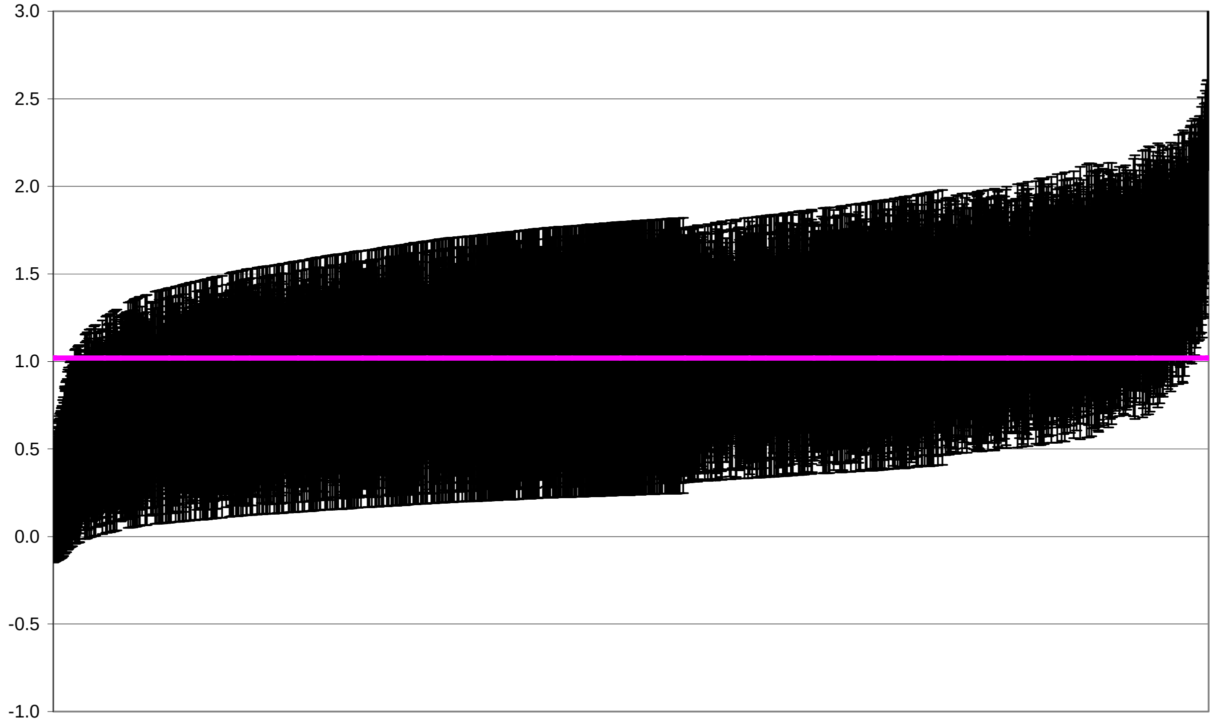


Figure 4-7. Mortality for Selected Conditions, Equal Weight

Figure 4-8. Mortality for Selected Conditions, Numerator Weight

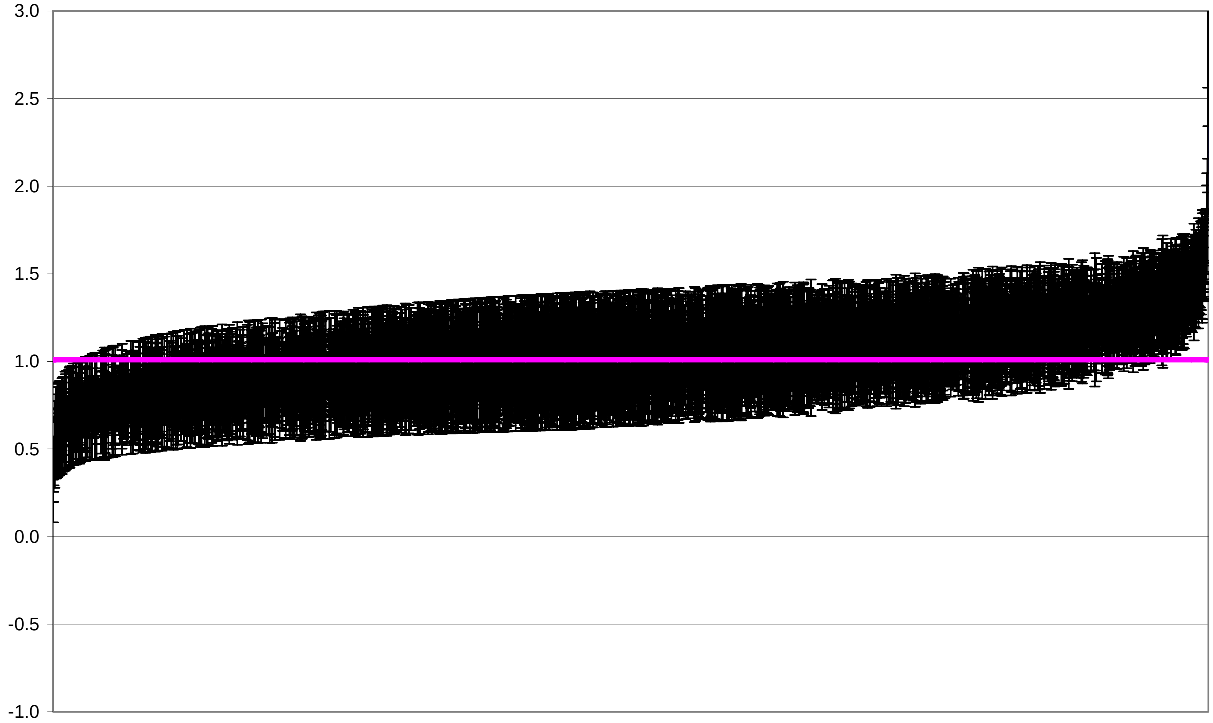
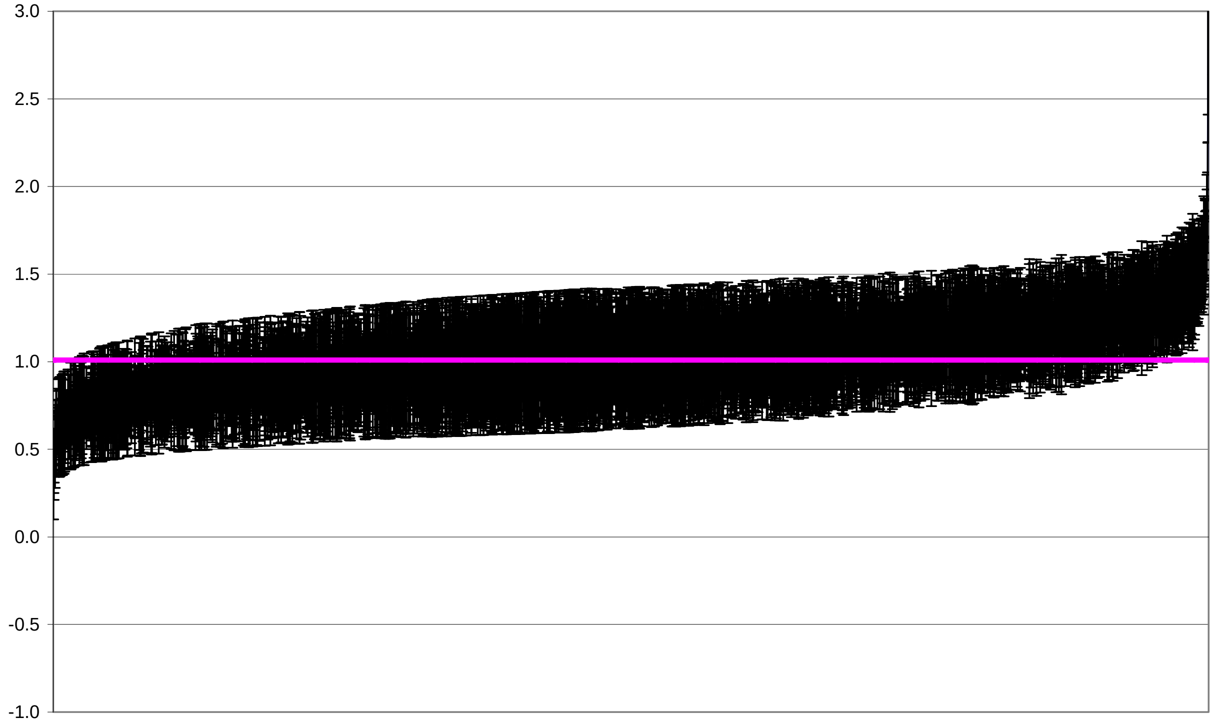


Figure 4-9. Mortality for Selected Conditions, Denominator Weight



Appendix D. Empirical Methods

D.1. 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_i = number of patients in sample at hospital i.
α = average outcome in the entire sample[6] (e.g. Ybar).

D.2 Computing the Noise Variance

Estimate the Risk Adjusted Ratio (RAR) and Noise Variance using the Ratio Method[7] of Indirect Standardization for each Hospital:

D.2.1. 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.

D.2.2. Estimating Variance of RAR (SE is the square root):

$Var(RAR_i)$
$= Var[\alpha(O_i/E_i)]$
$= (\alpha/E_i)^2 Var[O_i]$ (since $var(aX) = a^2 var(X)$ for any constant a)
$= (\alpha/E_i)^2 Var[(1/n_i)\sum(Y_{ij})]$ (by the definition of O_i)
$= (\alpha/E_i)^2 (1/n_i)^2 Var[\sum(Y_{ij})]$ (since $var(aX) = a^2 var(X)$ for any constant a)
$= (\alpha/E_i)^2 (1/n_i)^2 [\sum Var(Y_{ij})]$ (since $var(\sum X_i)=\sum var(X_i)$ if Xi are independent)
$= (\alpha/E_i)^2 (1/n_i)^2 \sum [P_{ij}(1-P_{ij})]$ (since Y is 0/1, so $var(Y) = P(1-P)$)

---

[6] For the AHRQ QI, the sample is the entire reference population consisting of the discharges in the SID for the participating states pooled over three years (2001-2003). Therefore, the "average outcome for the entire sample" is the population rate.
[7] Risk-adjusted rate = (Observed rate / Expected Rate) * Population Rate

D.3. Computing the Composite Variance

1) Setup[8]
   a) 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:
      i) $M = \mu + \varepsilon$
      ii) Let the KxK signal variance-covariance be $Var(\mu) = \Omega_\mu$
      iii) Let the KxK noise variance-covariance be $Var(\varepsilon) = \Omega_\varepsilon$
   b) Let $\hat{\mu}$ (1xK) be the posterior (filtered) estimate of μ, so that:
      i) $\mu = \hat{\mu} + \nu$, where the 1xK vector ν represents the prediction error of the posterior estimates, and Var(ν) is the KxK variance-covariance matrix for these posterior estimates.
   c) 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:
      i) $Var(\nu w) = w' Var(\nu) w$
         Thus, we simply need an estimate of Var(ν).

2) 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 based on different sample of patients or independent patient outcomes).
   a) Forming each measure in isolation, using superscripts to indicate the measure (k=1,…,K) as above, so:
      i) $\hat{\mu}^k = M^k \hat{\beta}^k = M^k \left[\Omega_\mu^{kk} + \Omega_\varepsilon^{kk}\right]^{-1} \Omega_\mu^{kk}$
      ii) $Var(\nu^k) = \Omega_\mu^{kk} - \Omega_\mu^{kk}\left(\Omega_\mu^{kk} + \Omega_\varepsilon^{kk}\right)^{-1}\Omega_\mu^{kk} = \Omega_\mu^{kk}\left(1 - \hat{\beta}^k\right)$
      iii) Note that in this simple case the filtered estimate is a simple shrinkage estimator and:
         (1) $\hat{\beta}^k$ is the signal ratio of measure k, also 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.
         (2) The variance of the filtered estimate is simply the signal variance times one minus the signal ratio. Thus, if the signal ratio is zero (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 zero if there is a signal ratio of 1 – no noise).
   b) 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, get the covariance elements, which are (for j≠k):
      i) $Cov(\nu^j, \nu^k) = E\left[\left(\mu^j - \hat{\mu}^j\right)\left(\mu^k - \hat{\mu}^k\right)\right]$

---

[8] For more information on the empirical bayes estimator methods, 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. Medical Care. 44(8):774-778, August 2006; and McClellan MB and Staiger DO, The Quality of Healthcare Providers, NBER Working Paper #7327, September, 1999 (at http://www.nber.org/papers/w7327).

ii) After some algebra (assuming independent estimation error in the two measures), one gets the following simple expression:
(1) $Cov\left(\nu^j, \nu^k\right) = \Omega_\mu^{jk}\left(1-\hat{\beta}^j\right)\left(1-\hat{\beta}^k\right)$

iii) Note that this is just the signal covariance, times one minus the signal ratio for each of the measures. Thus, if the signal ratio is zero 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 zero.

iv) Also note that if one measure is missing, then the signal ratio is simply set to zero – the filtered estimate is shrunk all the way back to the (conditional) mean, and the variance and covariance are as defined above.