AHRQ Quality Indicators

Patient Safety 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)

Patient Safety Indicators (PSI) Composite Measure Workgroup Final Report March 2008

Table of Contents

1. Introduction 1
2. Reasons for Composite Measures 1
3. Alternative Perspectives on Composite Measures 3
4. Methodology for the AHRQ QI Composite Measures 3
5. Performance of the AHRQ QI Composite Measures 10
6. Concluding Comments 11
Appendix A. AHRQ QI Composite Measure Workgroup A-1
Appendix B. IQI Composite Tables B-1
Appendix C. IQI Composite Figures C-1
Appendix D. Empirical Methods D-1

AHRQ Quality Indicators
Patient Safety Indicators
Composite Measure Workgroup
Final Report

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] 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 a composite measure for the Patient Safety Indicators (PSIs): *Patient Safety for Selected Indicators*.

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 Patient Safety 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 Patient Safety Indicators, including the selection criteria, coding, and specifications, see the Guide to Patient Safety Indicators and the Patient Safety 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 20 provider-level PSIs for various types of adverse events, making it difficult to formulate general statements about overall trends or differences in quality and patient safety.

[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/psi/psi_guide_v31.pdf; Technical Specifications: http://www.qualityindicators.ahrq.gov/downloads/psi/psi_technical_specs_v32.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 and patient safety.

- *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 improving patient safety.

- *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 and patient safety.

- *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 a single composite measure for the PSIs: *Patient Safety for Selected Indicators*. Appendix B presents PSI composite tables (Tables 1-8). Table 1 shows the reference population, including the incidence rate for each adverse event.

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 PSIs, one could advocate for separate composites based on the type of adverse event (e.g., postoperative). However, in general, the component indicators apply to the same providers (see Table 2) and show at least some positive correlation with one another (see Table 3). Therefore, the initial composite includes all the provider-level (nonobstetric) indicators (see table below).[4] Future development might examine subcomposites for certain indicators.

AHRQ PSI Composite Measure

Patient Safety for Selected Indicators	
PSI #03 Decubitus Ulcer	PSI #11 Postop Respiratory Failure
PSI #06 Iatrogenic Pneumothorax	PSI #12 Postop PE Or DVT
PSI #07 Selected Infection Due to Medical Care	PSI #13 Postop Sepsis
PSI #08 Postop Hip Fracture	PSI #14 Postop Wound Dehiscence
PSI #09 Postop Hemorrhage or Hematoma	PSI #15 Accidental Puncture or Laceration
PSI #10 Postop Physio and Metabol Derangmts	

Note: PE=pulmonary embolism; DVT=deep vein thrombosis.

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 the 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[5] 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

[4] Complications of Anesthesia (PSI #01) is not included because of the reliance on E-codes. Failure to Rescue (PSI #04) is not included because that indicator is already a composite. Low-Mortality DRGs (PSI #02), Foreign Body Left During Procedure (PSI #05), and Transfusion Reaction (PSI #16) are serious reportable events (i.e., "never events") and are reported as counts.
[5] 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.

population used in this analysis includes the States participating in the Healthcare Cost & Utilization Project (HCUP) for 2001-2003, consisting of 38 States and approximately 90 million discharges.[6]

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

Table 1 shows the reference population numerator, denominator, and rate for each PSI. 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 first 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).

RAR = [risk-adjusted ratio × weight] + [reference population ratio × (1 – weight)]

Table 4 shows the average reliability weights for the PSIs based on denominator size. For small providers, the weight is closer to 0. For large providers, the weight is closer to 1. For a given provider, if the denominator is 0, then the weight assigned is 0 (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 5 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. The reference population rate is the same among all providers (see Figures 1.1 and 1.2 in Appendix C).

[6] 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 - GHA: 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.

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., 1/11 = 0.0909).

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 a potentially preventable adverse event. One might also use weights that reflect the amount of excess mortality or complications associated with the adverse event, or the amount of confidence one has in identifying events (i.e., the positive predictive value).

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 5 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 PSI. 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

PSI	Average Annual Denominator	Observed Rate	Risk-Adjusted Rate	Rate Std. Error
PSI #03 Decubitus Ulcer	2,112	15.943	26.279	2.359
PSI #06 Iatrogenic Pneumothorax	5,995	2.224	1.619	0.153
PSI #07 Selected Infection Due to Medical Care	3,630	4.592	3.169	0.360
PSI #08 Postop Hip Fracture	2,248	0.297	0.509	0.262
PSI #09 Postop Hemorrhage or Hematoma	3,476	2.589	2.713	0.466
PSI #10 Postop Physio and Metabol Derangmts	3,233	0.722	0.846	0.343
PSI #11 Postop Respiratory Failure	3,044	9.417	6.802	0.817
PSI #12 Postop PE Or DVT	3,478	14.089	11.960	0.856
PSI #13 Postop Sepsis	764	18.333	17.123	1.997
PSI #14 Postop Wound Dehiscence	736	0.905	0.961	0.998
PSI #15 Accidental Puncture or Laceration	6,285	8.697	4.869	0.324

Note: Observed and risk-adjusted rate are per 1,000. PE=pulmonary embolism; DVT=deep vein thrombosis.

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

PSI	Reference Population Rate	Risk-Adjusted Ratio	Ratio Std. Error
PSI #03 Decubitus Ulcer	22.081	1.190	0.107
PSI #06 Iatrogenic Pneumothorax	0.582	2.784	0.263
PSI #07 Selected Infection Due to Medical Care	2.054	1.543	0.175
PSI #08 Postop Hip Fracture	0.272	1.868	0.964
PSI #09 Postop Hemorrhage or Hematoma	2.175	1.247	0.215
PSI #10 Postop Physio and Metabol Derangmts	0.986	0.859	0.348
PSI #11 Postop Respiratory Failure	8.802	0.773	0.093
PSI #12 Postop PE Or DVT	9.169	1.304	0.093
PSI #13 Postop Sepsis	10.009	1.711	0.200
PSI #14 Postop Wound Dehiscence	2.080	0.462	0.480
PSI #15 Accidental Puncture or Laceration	3.612	1.348	0.090

Step 3. Compute the reliability-adjusted ratio

Step 3A. Compute the reliability weight

PSI	Ratio Std. Error	Noise Variance	Signal Variance	Reliability Weight
PSI #03 Decubitus Ulcer	0.107	0.0114	0.2208	0.9509
PSI #06 Iatrogenic Pneumothorax	0.263	0.0694	0.2295	0.7679
PSI #07 Selected Infection Due to Medical Care	0.175	0.0307	0.2854	0.9030
PSI #08 Postop Hip Fracture	0.964	0.9288	0.0892	0.0876
PSI #09 Postop Hemorrhage or Hematoma	0.215	0.0460	0.1320	0.7415
PSI #10 Postop Physio and Metabol Derangmts	0.348	0.1211	0.2940	0.7083
PSI #11 Postop Respiratory Failure	0.093	0.0086	0.2082	0.9603
PSI #12 Postop PE Or DVT	0.093	0.0087	0.2542	0.9668
PSI #13 Postop Sepsis	0.200	0.0398	0.1579	0.7986
PSI #14 Postop Wound Dehiscence	0.480	0.2302	0.2226	0.4916
PSI #15 Accidental Puncture or Laceration	0.090	0.0080	0.2247	0.9655

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 3B. Compute the reliability-adjusted ratio

PSI	Reliability Weight	Risk-Adjusted Ratio	Reference Population Ratio	Reliability-Adjusted Ratio
PSI #03 Decubitus Ulcer	0.9509	1.190	0.983	1.180
PSI #06 Iatrogenic Pneumothorax	0.7679	2.784	0.963	2.361
PSI #07 Selected Infection Due to Medical Care	0.9030	1.543	0.938	1.484
PSI #08 Postop Hip Fracture	0.0876	1.868	1.020	1.094
PSI #09 Postop Hemorrhage or Hematoma	0.7415	1.247	1.003	1.184
PSI #10 Postop Physio and Metabol Derangmts	0.7083	0.859	0.910	0.874
PSI #11 Postop Respiratory Failure	0.9603	0.773	0.965	0.781
PSI #12 Postop PE Or DVT	0.9668	1.304	0.982	1.293
PSI #13 Postop Sepsis	0.7986	1.711	0.936	1.555
PSI #14 Postop Wound Dehiscence	0.4916	0.462	1.004	0.738
PSI #15 Accidental Puncture or Laceration	0.9655	1.348	0.926	1.333

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 denominator weight.

	Denominator Weight
PSI #03 Decubitus Ulcer	0.0755
PSI #06 Iatrogenic Pneumothorax	0.2246
PSI #07 Selected Infection Due to Medical Care	0.1864
PSI #08 Postop Hip Fracture	0.0473
PSI #09 Postop Hemorrhage or Hematoma	0.0712
PSI #10 Postop Physio and Metabol Derangmts	0.0344
PSI #11 Postop Respiratory Failure	0.0280
PSI #12 Postop PE Or DVT	0.0709
PSI #13 Postop Sepsis	0.0086
PSI #14 Postop Wound Dehiscence	0.0152
PSI #15 Accidental Puncture or Laceration	0.2378

Step 5. Construct the composite measure

	Denominator Weight (A)	Reliability-Adjusted Ratio (B)	(A) × (B)
PSI #03 Decubitus Ulcer	0.0755	1.180	0.089
PSI #06 Iatrogenic Pneumothorax	0.2246	2.361	0.530
PSI #07 Selected Infection Due to Medical Care	0.1864	1.484	0.277
PSI #08 Postop Hip Fracture	0.0473	1.094	0.052
PSI #09 Postop Hemorrhage or Hematoma	0.0712	1.184	0.084
PSI #10 Postop Physio and Metabol Derangmts	0.0344	0.874	0.030
PSI #11 Postop Respiratory Failure	0.0280	0.781	0.022
PSI #12 Postop PE Or DVT	0.0709	1.293	0.092
PSI #13 Postop Sepsis	0.0086	1.555	0.013
PSI #14 Postop Wound Dehiscence	0.0152	0.738	0.011
PSI #15 Accidental Puncture or Laceration	0.2378	1.333	0.317
Patient Safety for Selected Indicators			*1.517*
Standard Error			*0.072*
Confidence Interval at p<0.05		*1.376*	*1.659*

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 iatrogenic pneumothorax 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 iatrogenic pneumothorax as the best mechanism to lower the composite score.

5. Performance of the AHRQ QI Composite Measures

5.1. Evaluation Criteria

Tables 6-8 in Appendix B and Figures 2.1-2.5 and 3.1-3.5 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 and patient safety or other factors that might influence quality and patient safety.

5.2. Results

Table 6 shows the discrimination performance of the composite measure *Patient Safety for Selected Indicators*. 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. The single and equal weights have the least ability to discriminate. The single indicator used as an example is "selected infection due to medical care." The numerator weight tends to have the greatest ability to discriminate, followed by the denominator weight and factor weight.

In general, the composite identifies a large number of providers with performance that is better or worse than average. Figures 2.1-2.5 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.

Table 7 shows the forecasting performance of the composite measure. 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 8 shows the correlation among the composite measures using the alternative weights. For the *Patient Safety for Selected Indicators*, the correlations range from 0.529 to 0.979. Regardless of the weight used, 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 or patient safety 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. PSI Composite Tables

Table 1. Reference Population

PSI #03 Decubitus Ulcer	476,583	21,583,071	22.081
PSI #06 Iatrogenic Pneumothorax	37,335	64,193,131	0.582
PSI #07 Selected Infection Due to Medical Care	109,442	53,292,737	2.054
PSI #08 Postop Hip Fracture	3,685	13,533,878	0.272
PSI #09 Postop Hemorrhage or Hematoma	44,250	20,347,679	2.175
PSI #10 Postop Physio and Metabol Derangmts	9,700	9,841,216	0.986
PSI #11 Postop Respiratory Failure	70,440	8,002,305	8.802
PSI #12 Postop PE Or DVT	185,794	20,263,685	9.169
PSI #13 Postop Sepsis	24,633	2,461,073	10.009
PSI #14 Postop Wound Dehiscence	9,038	4,346,106	2.080
PSI #15 Accidental Puncture or Laceration	245,532	67,971,505	3.612

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

Table 2. Provider-level Rates

		Risk Adjusted		Reliability Adjusted	
PSI	**Hospitals**	**Rate**	**Std. Dev.**	**Rate**	**Std. Dev.**
PSI #03 Decubitus Ulcer	4,823	19.856	19.767	20.368	9.963
PSI #06 Iatrogenic Pneumothorax	4,909	0.413	0.636	0.540	0.175
PSI #07 Selected Infection Due to Medical Care	4,908	1.528	2.558	1.768	1.191
PSI #08 Postop Hip Fracture	4,312	0.419	8.297	0.278	0.021
PSI #09 Postop Hemorrhage or Hematoma	4,356	1.860	3.701	2.146	0.474
PSI #10 Postop Physio and Metabol Derangmts	3,603	0.967	10.712	0.898	0.293
PSI #11 Postop Respiratory Failure	3,592	7.672	11.257	8.404	3.687
PSI #12 Postop PE Or DVT	4,352	7.852	12.055	8.248	3.368
PSI #13 Postop Sepsis	3,398	9.727	14.253	9.597	2.958
PSI #14 Postop Wound Dehiscence	4,004	1.984	3.377	2.099	0.528
PSI #15 Accidental Puncture or Laceration	4,909	2.720	3.372	3.205	1.363

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

Table 3. Provider-Level Correlation

PSI	PSI #03	PSI #06	PSI #07	PSI #08	PSI #09	PSI #10
PSI #03 Decubitus Ulcer	1.000	0.106	0.198	0.034	-0.027	0.060
PSI #06 Iatrogenic Pneumothorax		1.000	0.404	0.003	0.194	0.042
PSI #07 Selected Infection Due to Medical Care			1.000	0.005	0.174	0.082
PSI #08 Postop Hip Fracture				1.000	-0.005	0.000
PSI #09 Postop Hemorrhage or Hematoma					1.000	0.016
PSI #10 Postop Physio and Metabol Derangmts						1.000
PSI #11 Postop Respiratory Failure	0.155	0.056	0.116	-0.001	0.024	0.131
PSI #12 Postop PE Or DVT	0.192	0.121	0.253	-0.001	0.034	0.081
PSI #13 Postop Sepsis	0.171	0.033	0.161	0.016	-0.010	0.055
PSI #14 Postop Wound Dehiscence	0.053	0.085	0.005	0.013	0.071	0.035
PSI #15 Accidental Puncture or Laceration	-0.003	0.375	0.301	0.002	0.273	0.074
PSI	**PSI #11**	**PSI #12**	**PSI #13**	**PSI #14**	**PSI #15**	
PSI #11 Postop Respiratory Failure	1.000	0.122	0.221	0.054	-0.037	
PSI #12 Postop PE Or DVT		1.000	0.108	0.010	0.052	
PSI #13 Postop Sepsis			1.000	0.006	-0.061	
PSI #14 Postop Wound Dehiscence				1.000	0.054	
PSI #15 Accidental Puncture or Laceration					1.000	

Source: HCUP State Inpatient Data, 2001-2003.

Table 4. Reliability Weight by Average Annual Denominator

Average Annual Denominator Size (by quartile)					
PSI	**Hospitals**	**Q1**	**Q2**	**Q3**	**Q4**
PSI #03 Decubitus Ulcer	4,823	53.3	367.3	1,325.2	4,219.7
PSI #06 Iatrogenic Pneumothorax	4,909	276.6	1,248.1	3,939.3	11,965.4
PSI #07 Selected Infection Due to Medical Care	4,908	210.7	1,020.1	3,285.3	9,961.7
PSI #08 Postop Hip Fracture	4,312	25.3	207.9	774.8	3,176.8
PSI #09 Postop Hemorrhage or Hematoma	4,356	30.5	305.8	1,181.1	4,710.8
PSI #10 Postop Physio and Metabol Derangmts	3,603	17.2	162.4	650.2	2,811.1
PSI #11 Postop Respiratory Failure	3,592	15.9	150.1	574.4	2,230.1
PSI #12 Postop PE Or DVT	4,352	30.7	306.9	1,177.4	4,693.2
PSI #13 Postop Sepsis	3,398	6.6	46.0	164.1	748.5
PSI #14 Postop Wound Dehiscence	4,004	18.1	112.9	324.0	992.3
PSI #15 Accidental Puncture or Laceration	4,909	285.9	1,284.4	4,093.2	12,791.5

Average Reliability Weight					
PSI	**Q1**	**Q2**	**Q3**	**Q4**	**Weighted Average**
PSI #03 Decubitus Ulcer	0.4062	0.7970	0.9253	0.9807	0.9584
PSI #06 Iatrogenic Pneumothorax	0.0861	0.2641	0.5638	0.8101	0.7288
PSI #07 Selected Infection Due to Medical Care	0.2536	0.5670	0.8215	0.9367	0.8872
PSI #08 Postop Hip Fracture	0.0043	0.0155	0.0542	0.1714	0.1741
PSI #09 Postop Hemorrhage or Hematoma	0.0265	0.1894	0.4826	0.7650	0.7202
PSI #10 Postop Physio and Metabol Derangmts	0.0237	0.1015	0.3083	0.6554	0.6263
PSI #11 Postop Respiratory Failure	0.1212	0.4432	0.7361	0.9040	0.8664
PSI #12 Postop PE Or DVT	0.1776	0.6277	0.8797	0.9624	0.9381
PSI #13 Postop Sepsis	0.0575	0.2168	0.4372	0.7320	0.7023
PSI #14 Postop Wound Dehiscence	0.0248	0.1380	0.3114	0.5436	0.4904
PSI #15 Accidental Puncture or Laceration	0.1905	0.5571	0.8557	0.9614	0.9099

Source: HCUP State Inpatient Data, 2001-2003.

Table 5. Alternative Composite Weights

PSI	Single Indicator Weight	Equal Weight	Numerator Weight	Denominator Weight	Factor Weight
PSI #03 Decubitus Ulcer	0.0000	0.0909	**0.3918**	0.0755	0.1017
PSI #06 Iatrogenic Pneumothorax	0.0000	0.0909	0.0307	0.2246	0.0939
PSI #07 Selected Infection Due to Medical Care	**1.0000**	0.0909	0.0900	0.1864	**0.1470**
PSI #08 Postop Hip Fracture	0.0000	0.0909	0.0030	0.0473	0.0252
PSI #09 Postop Hemorrhage or Hematoma	0.0000	0.0909	0.0364	0.0712	0.0633
PSI #10 Postop Physio and Metabol Derangmts	0.0000	0.0909	0.0080	0.0344	0.1145
PSI #11 Postop Respiratory Failure	0.0000	0.0909	0.0579	0.0280	0.1151
PSI #12 Postop PE Or DVT	0.0000	0.0909	0.1527	0.0709	0.1200
PSI #13 Postop Sepsis	0.0000	0.0909	0.0203	0.0086	0.1252
PSI #14 Postop Wound Dehiscence	0.0000	0.0909	0.0074	0.0152	0.0337
PSI #15 Accidental Puncture or Laceration	0.0000	0.0909	0.2018	**0.2378**	0.0604

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

Table 6. Discrimination Performance of Alternative Composites

Composite	Providers	Better Than Average	Average	Worse Than Average
Patient Safety for Selected Indicators				
Single Indicator Weight	4,901	7.37%	85.86%	6.77%
Equal Weight	4,901	10.20%	81.45%	8.35%
Numerator Weight	4,902	28.34%	57.81%	13.85%
Denominator Weight	4,908	17.77%	70.60%	11.63%
Factor Weight	4,910	16.74%	73.46%	9.80%

Source: HCUP State Inpatient Data, 2001-2003.

Table 7. Forecasting Performance of Alternative Composites

PSI	PSI #03	PSI #06	PSI #07	PSI #08	PSI #09	PSI #10
Patient Safety for Selected Indicators						
Single Indicator Weight						
Best 20%	-0.070*	-0.161*	**-0.488***	0.001	-0.076*	-0.076*
Worst 20%	0.148*	0.281*	**0.864***	0.017*	0.096*	0.154*
Equal Weight						
Best 20%	-0.263*	-0.220*	-0.344*	-0.015*	-0.143*	-0.153*
Worst 20%	0.252*	0.385*	0.617*	0.024*	0.198*	0.389*
Numerator Weight						
Best 20%	**-0.364***	-0.150*	-0.367*	-0.009*	-0.090*	-0.077*
Worst 20%	**0.590***	0.271*	0.431*	0.017*	0.062*	0.195*
Denominator Weight						
Best 20%	-0.174*	-0.291*	-0.408*	-0.003	-0.130*	-0.066*
Worst 20%	0.177*	0.493*	0.663*	0.022*	0.190*	0.229*
Factor Weight						
Best 20%	-0.278*	-0.199*	**-0.372***	-0.016*	-0.109*	-0.169*
Worst 20%	0.263*	0.376*	**0.644***	0.020*	0.172*	0.396*

PSI	PSI #11	PSI #12	PSI #13	PSI #14	PSI #15
Patient Safety for Selected Indicators					
Single Indicator Weight					
Best 20%	-0.057*	-0.130*	-0.054*	0.020	-0.186*
Worst 20%	0.082*	0.205*	0.110*	0.036*	0.221*
Equal Weight					
Best 20%	-0.222*	-0.190*	-0.149*	-0.092*	-0.232*
Worst 20%	0.305*	0.380*	0.187*	0.153*	0.338*
Numerator Weight					
Best 20%	-0.141*	-0.186*	-0.086*	-0.024*	-0.287*
Worst 20%	0.148*	0.412*	0.100*	0.054*	0.274*
Denominator Weight					
Best 20%	-0.068*	-0.142*	-0.053*	-0.029*	**-0.354***
Worst 20%	0.107*	0.256*	0.052*	0.061*	**0.526***
Factor Weight					
Best 20%	-0.256*	-0.207*	-0.168*	-0.040*	-0.151*
Worst 20%	0.316*	0.408*	0.206*	0.099*	0.275*

Source: HCUP State Inpatient Data, 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 8. Correlation of Alternative Composites

Composite	Single Indicator Weight	Equal Weight	Numerator Weight	Denominator Weight	Factor Weight
Patient Safety for Selected Indicators					
Single Indicator Weight	1.000	0.655	0.529	0.753	0.715
Equal Weight		1.000	0.815	0.866	0.979
Numerator Weight			1.000	0.742	0.807
Denominator Weight				1.000	0.828
Factor Weight					1.000

Source: HCUP State Inpatient Data, 2001-2003.

Appendix C. Composite Figures

1. Single Indicator Composites

Figure 1.1 – PSI #7 Selected Infection Due to Medical Care

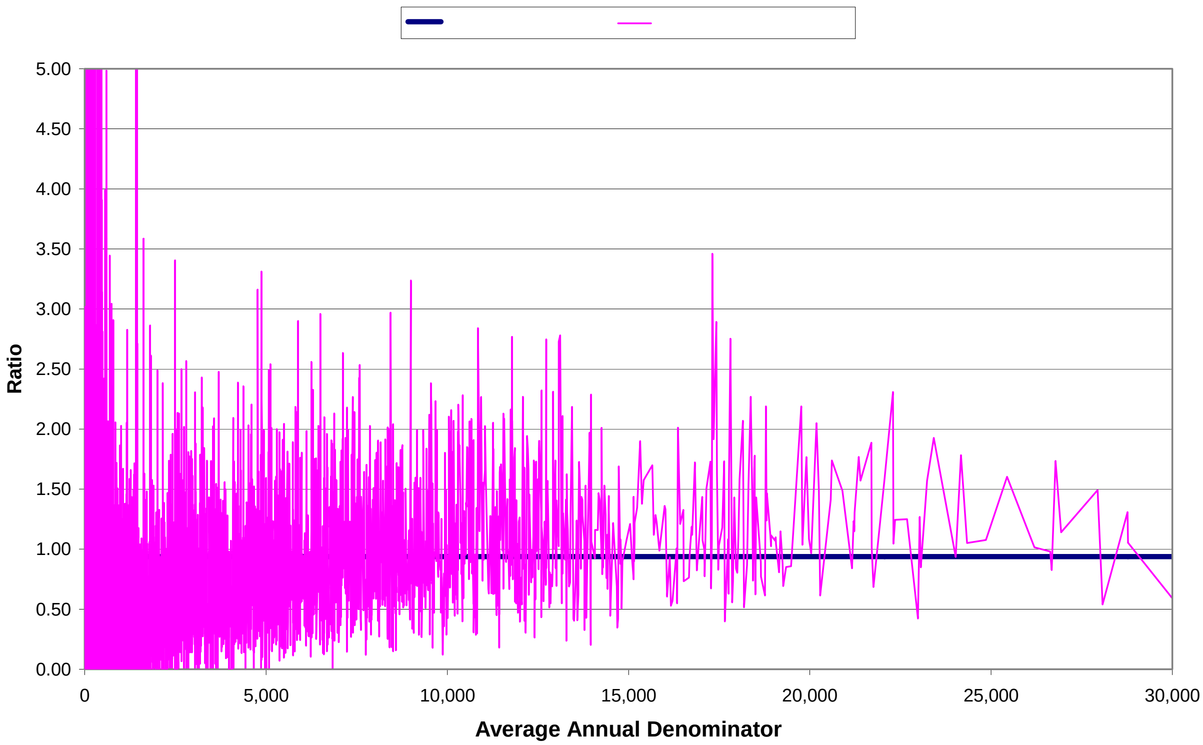
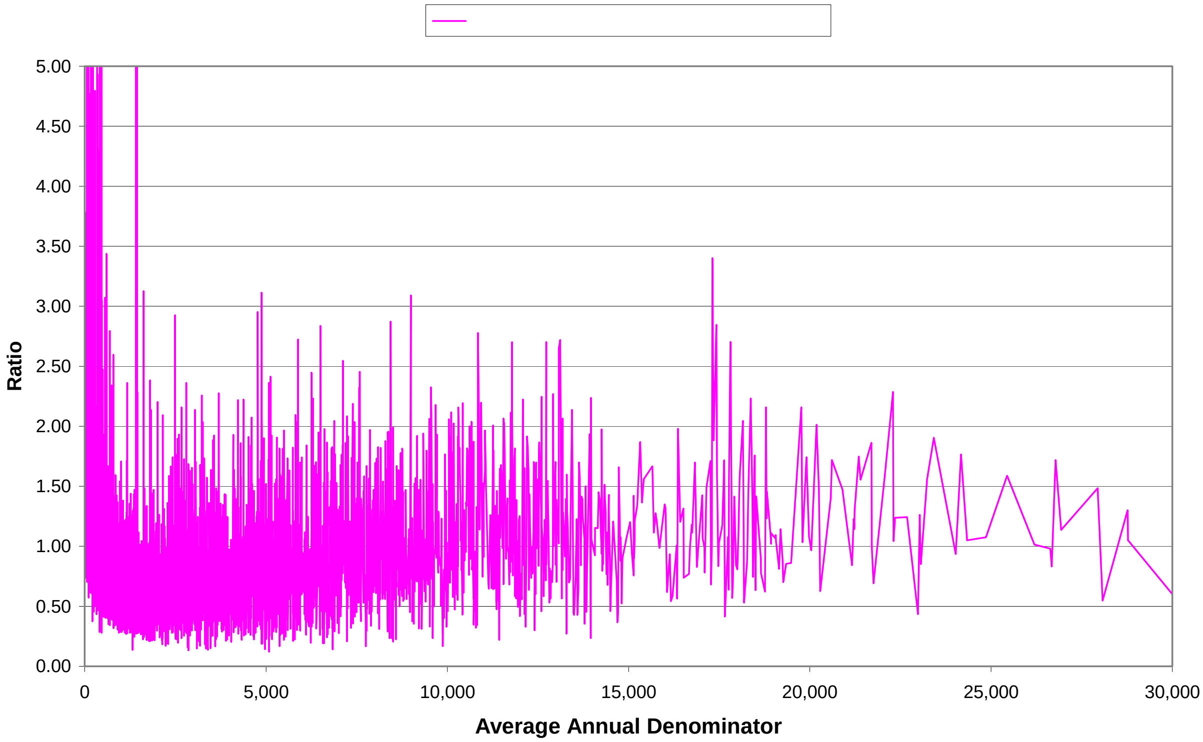


Figure 1.2 – PSI #7 Selected Infection Due to Medical Care



2. Precision of Alternative Composites

Figure 2.1 - Patient Safety for Selected Indicators, Single Indicator Weight

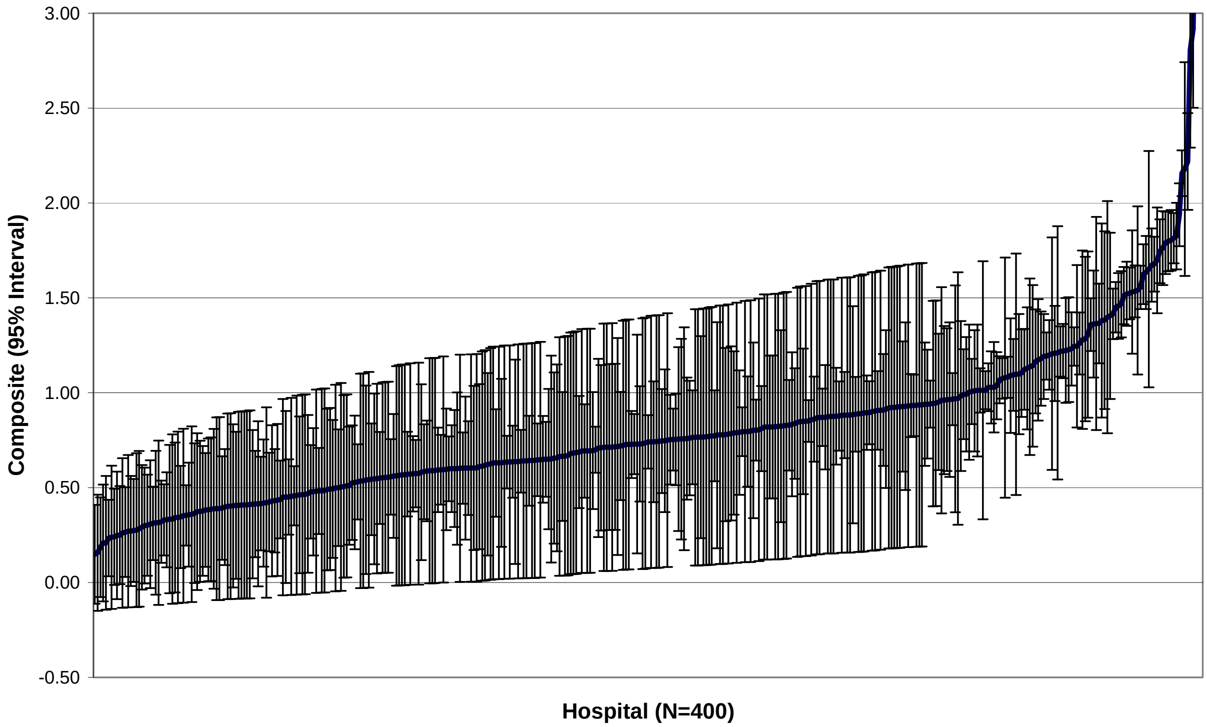


Figure 2.2 - Patient Safety for Selected Indicators, Equal Weight

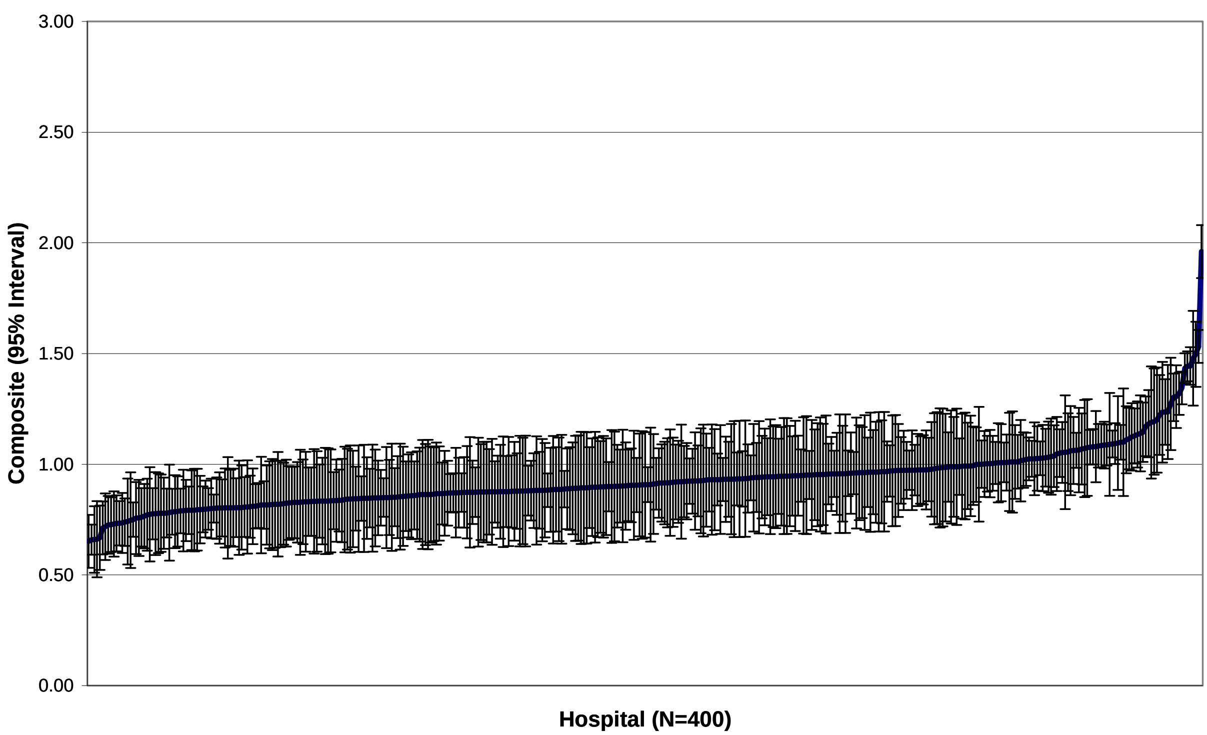


Figure 2.3 - Patient Safety for Selected Indicators, Numerator Weight

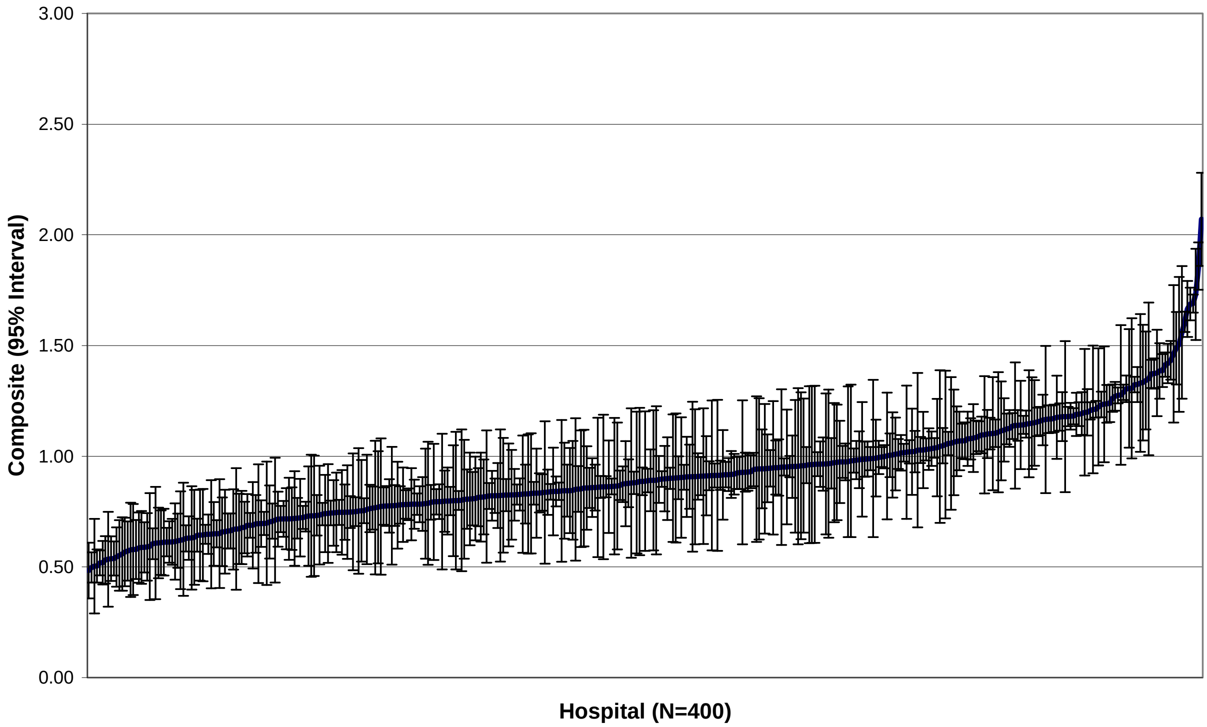


Figure 2.4 - Patient Safety for Selected Indicators, Denominator Weight

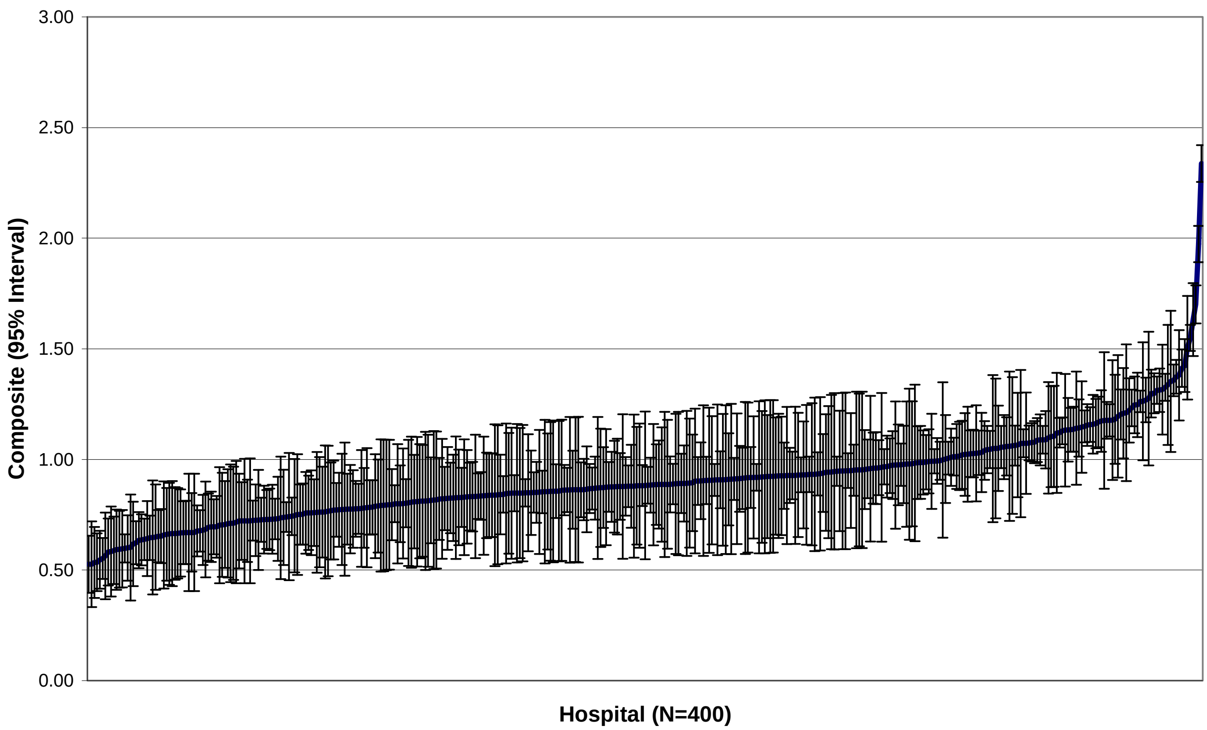
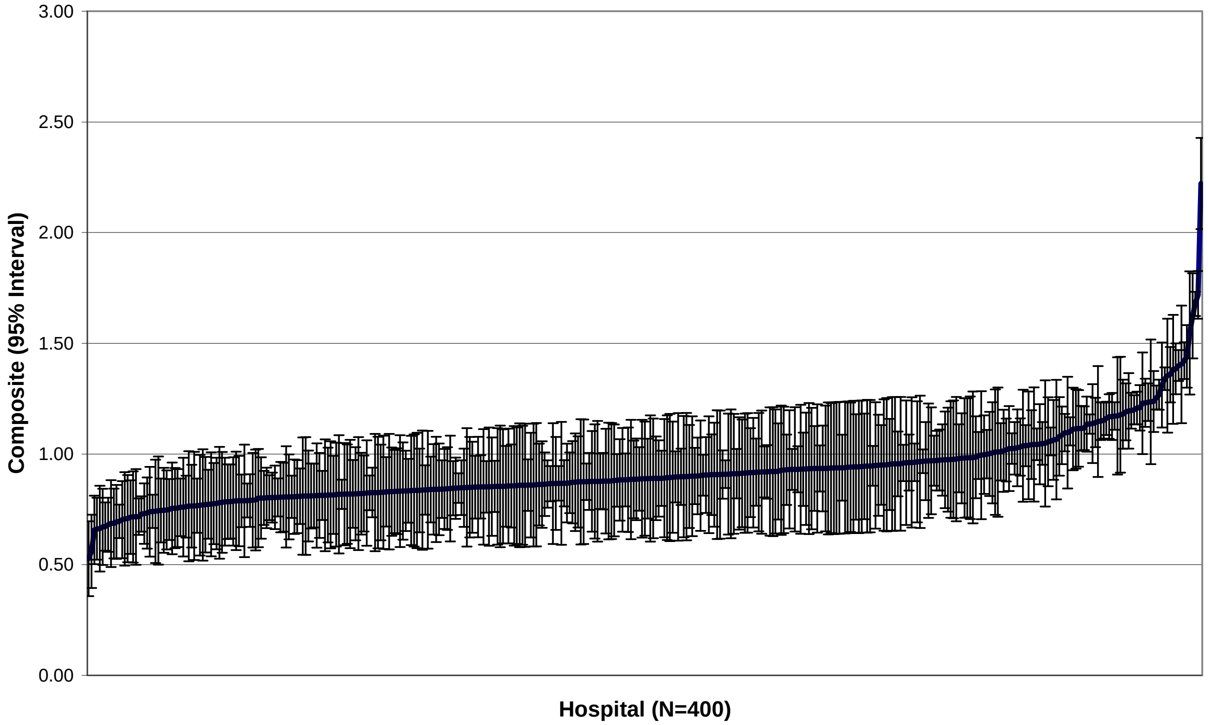


Figure 2.5 - Patient Safety for Selected Indicators, Factor Weight



3. *Distribution of Alternative Composites*

Figure 3.1 - Patient Safety for Selected Indicators, Single Indicator Weight

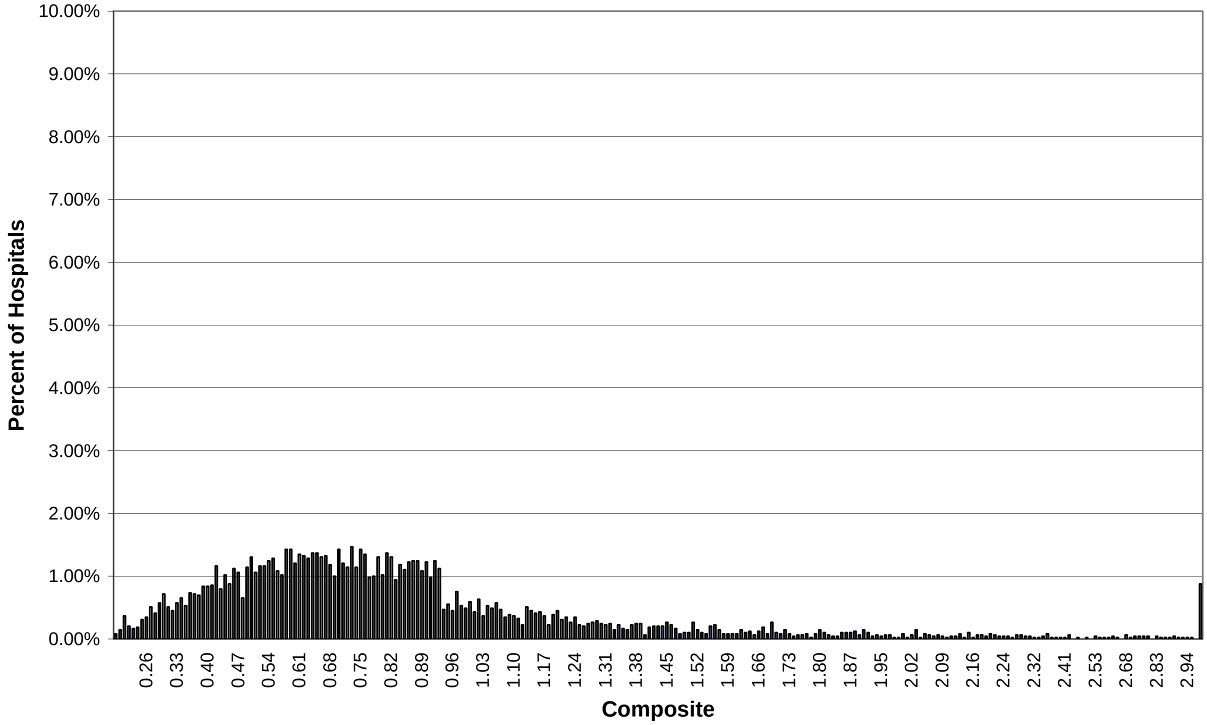


Figure 3.2 - Patient Safety for Selected Indicators, Equal Weight

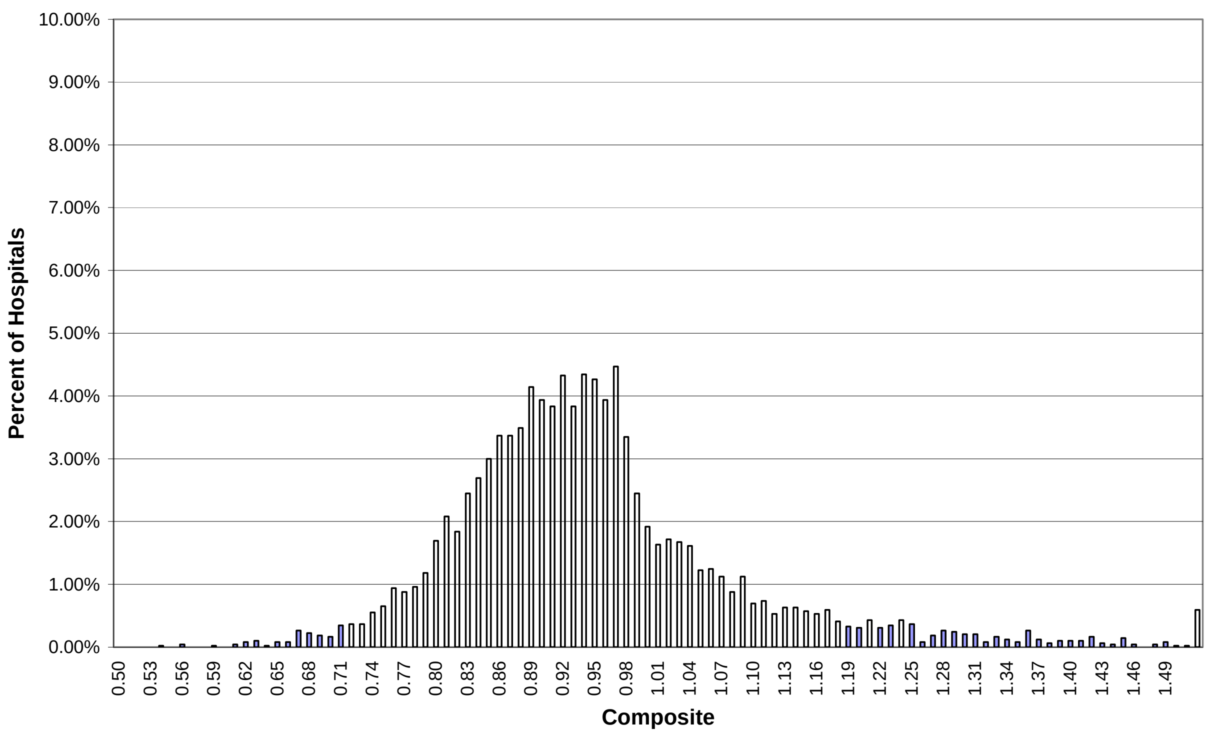


Figure 3.3 - Patient Safety for Selected Indicators, Numerator Weight

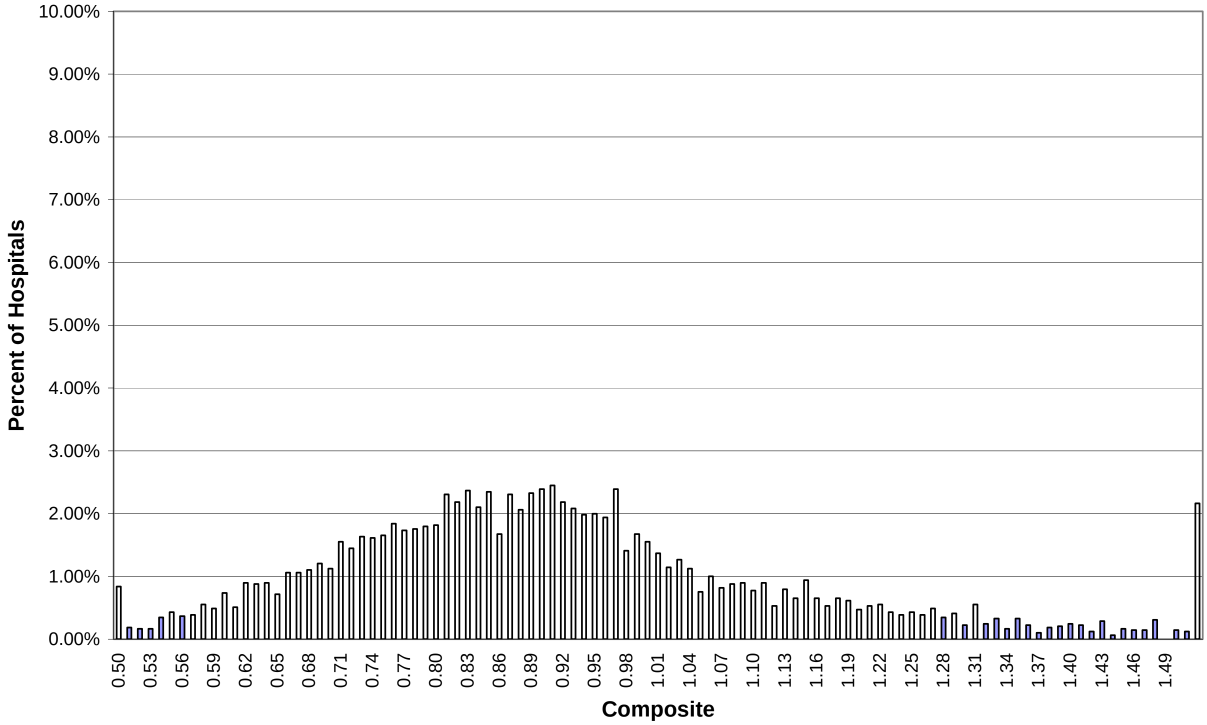


Figure 3.4 - Patient Safety for Selected Indicators, Denominator Weight

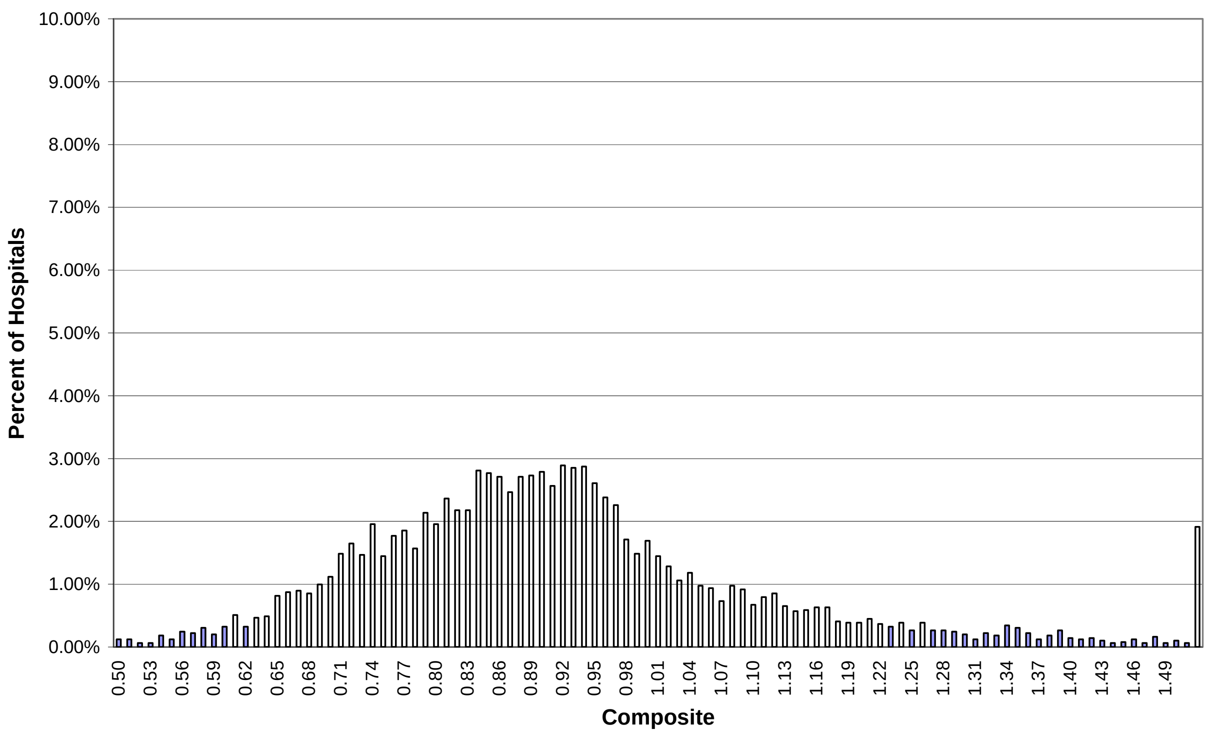
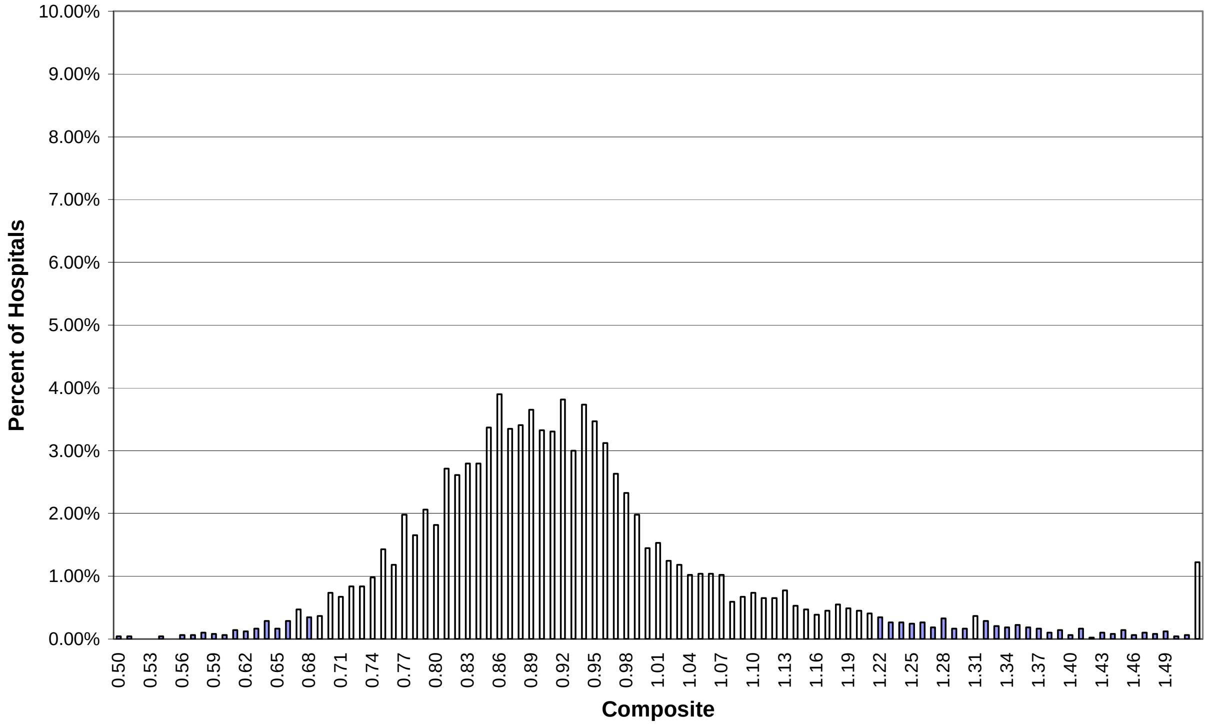


Figure 3.5 - Patient Safety for Selected Indicators, Factor Weight



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_i = 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 (SE is the square root of the variance)

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

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 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. Med Care 2006 Aug;44(8):774-8; and McClellan M and Staiger D, 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, so:

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

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

   - 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 $Var\left(\nu^{k}\right)$ 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):

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

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

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

- 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.