

*AHRQ Quality Indicators*TM



Quality Indicator Empirical Methods

Prepared for:

Agency for Healthcare Research and Quality
U.S. Department of Health and Human Services
540 Gaither Road
Rockville, MD 20850
<http://www.qualityindicators.ahrq.gov>

Contract No. HHSA290201200003I

Updated by:

Truven Health Analytics
Stanford University (prime contractor)

Revised November 2015

Disclaimer

This document is based on research conducted by Stanford University and Truven Health Analytics under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. HHSA290201200003I). This document reflects an update of work products developed by Battelle under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. HHSA290201200001C). The findings and conclusions in this document are those of the author(s), who are responsible for its content, and do not necessarily represent the views of AHRQ. No statement in this report should be construed as an official position of AHRQ or of the U.S. Department of Health and Human Services.

The information in this report is intended to help clinicians, employers, policymakers, and others make informed decisions about the provision of health care services. This document is intended as a reference and not as a substitute for clinical judgment.

This document is in the public domain and may be used and reprinted without permission except those copyrighted materials that are clearly noted in the document. Further reproduction of those copyrighted materials is prohibited without the specific permission of copyright holders.

Suggested Citation:

Quality Indicator Empirical Methods. Revised by Truven Health Analytics, Stanford University (prime contractor), under Contract No. HHSA290201200003I. Rockville, MD: Agency for Healthcare Research and Quality; November 2014.

None of the investigators has any affiliations or financial involvement that conflicts with the material presented in this document.

Contents

Appendix Table A.1. AHRQ QI Risk-Adjustment and Uses of POA.....	6
Other Helpful Documents.....	8
Chapter 1. The User’s Dataset.....	9
Chapter 2. Calculating Volume and Count Indicators.....	11
Chapter 3. Calculating Area-Level Indicators—Observed Rates.....	13
Chapter 4. Risk Adjustment for Area-Level Indicators.....	16
Chapter 5. Calculating Area-Level Indicators—Expected, Risk-Adjusted, and Smoothed Rates....	17
Chapter 6. Overview of Provider-Level Quality Indicators and Present-on-Admission	19
Chapter 7. Calculating Provider-Level Observed Rates.....	22
Chapter 8. Risk Adjustment for Provider-Level Indicators.....	25
Chapter 9. Estimating Composite Measures	30
Chapter 10. Specifications for AHRQ QI Reference Population	38
Chapter 11. Estimating Risk Adjustment Models	39
Chapter 12. Other Files Referenced in AHRQ QI Software	42
Appendix A. Table of AHRQ QI Risk-Adjustment / POA	44
Appendix B. Table of AHRQ QI Provider-Level Risk-Adjustment Covariates	44
Appendix C. External Resources Referenced by the AHRQ QI.....	49

List of Tables

Table 1.1 Required Data Elements	9
Table 1.2 Data Elements and Data Values to Be Constructed by the User	10
Table 1.3 Analysis Data Inclusion Rule	10
Table 2.1 AHRQ QI Volume Indicators.....	11
Table 2.2 AHRQ QI Count Indicators.....	11
Table 3.1 AHRQ QIs Area-Level Indicators.....	13

Table 6.1 Methods Used by QI Software to Distinguish Complications From Comorbidities	20
Table 6.2 Values for the Present-on-Admission Data Element	21
Table 9.1 Weights of PSI-90 component indicators	34
Appendix Table A.1. AHRQ QI Risk-Adjustment and Uses of POA	44
Appendix Table B.1. AHRQ QI Risk-Adjustment Covariates for Provider-Level Indicators	47

Overview

This document describes the empirical methods used to calculate the Agency for Healthcare Research and Quality Quality Indicators™ (AHRQ QIs). The QIs measure health care quality and can be used to highlight potential quality concerns, identify areas that need further study and investigation, and track changes over time. The QIs are calculated using software that is freely available at www.qualityindicators.ahrq.gov.

The current AHRQ QI modules represent various aspects of quality:

- Prevention Quality Indicators (PQIs) identify hospital admissions that evidence suggests might have been avoided given access to high-quality health care, preventive care and health promoting resources within a community (first released November 2000, last updated November 2015).
- Inpatient Quality Indicators (IQIs) reflect quality of care inside hospitals, as well as across geographic areas, including inpatient mortality for medical conditions and surgical procedures (first released May 2002, last updated November 2015).
- Patient Safety Indicators (PSIs) reflect quality of care inside hospitals, as well as geographic areas, to focus on potentially avoidable complications and iatrogenic events (first released March 2003, last updated November 2015).
- Pediatric Quality Indicators (PDIs) use indicators from the other three modules with adaptations for use among children and neonates to reflect quality of care inside hospitals, as well as geographic areas and to identify potentially avoidable hospitalizations (first released April 2006, last updated November 2015). The Neonatal Quality Indicators (NQIs) reflect quality of care inside hospitals for at-risk neonates.

The AHRQ QI software is intended for use with discharge-level administrative records from inpatient hospital stays; this document often refers to them as *discharge records*. Each indicator can be described as giving results at either the *provider level* (i.e., Did the patient experience an adverse quality-related event while in the health care provider's facility?) or *area level* (i.e., Was the inpatient admission for a condition that might have been avoided if the patient's area of the country had more or better preventive or outpatient care?). As a practical matter, in the default configuration of the QI software, provider level is synonymous with hospital level and area level is synonymous with county level. Some indicators report the number of times that a hospital performed a medical procedure of interest. These volume indicators do not have denominators. Most of the AHRQ QIs are ratios in which the numerator is a count of hospitalizations with the condition or outcome of interest and the denominator is an estimate of the population (or hospitalizations) at risk for that outcome. The QI software calculates different kinds of rates:

- **Observed rate**—Conceptually, provider-level rates are the number of discharge records in which the patient experienced the QI adverse event divided by the number of discharge records at risk for the event; area-level rates are the number of hospitalizations for the condition of interest divided by the number of individuals who live in that area who are at risk for the condition.

- **Expected rate**—A comparative rate that incorporates information about an external reference population that is not part of the user’s input dataset—that is, the rate that would be predicted if the expected level of care observed in the reference population and estimated with risk adjustment regression models were applied to the mix of patients with demographic and comorbidity distributions observed in the user’s dataset. The expected rate is calculated only for risk-adjusted indicators. [Chapter 4](#) describes the QI reference population.
- **Risk-adjusted rate**—A comparative rate that also incorporates information about a reference population that is not part of the input dataset—that is, the rate that would be observed if the level of care observed in the user’s dataset were applied to a mix of patients with demographics and comorbidities distributed like the reference population. Appendix A lists the QIs that are risk adjusted.
- **Smoothed rate**—A weighted average of the risk-adjusted rate from the user’s input dataset and the rate observed in the reference population discharges; the smoothed rate is calculated with a shrinkage estimator (1) to result in a rate near that from the user’s dataset if the provider’s (or area’s) rate is estimated in a stable fashion with minimal noise or (2) to result in a rate near that of the reference population if the rate from the input dataset is unstable and based on noisy data. In practice, the smoothed rate brings rates toward the grand population mean (i.e., the rate among all discharges in the reference population) and does this more so for hospitals with lower volume (smaller denominators) and outliers (such as rural hospitals). Rates for larger, high volume, hospitals will tend not to move much with smoothing, even if their rate differs from the reference population rate.

In data collected beginning October 1, 2007, each diagnosis code may be accompanied by a data element that indicates whether the diagnosed condition was (1) Present-on-Admission (POA) and is therefore a pre-existing comorbidity or (2) developed during the hospitalization of interest and therefore is a complication. The handling of missing POA information in the AHRQ QI software has changed over time. Starting with version 6.0 of the QI software, POA data are assumed to be present on the input dataset. In version 5.0 and earlier, the user could select whether to compute risk-adjusted rates using the POA information or ignoring POA information. In versions of QI software prior to 5.0, a prediction module was used to impute missing POA information. Beginning with version 5.0, the prediction module has been removed and missing POA information is treated as if the condition is not POA.

In this document, we begin with a brief description of the dataset that a user must assemble to run the QI software and then describe the methods associated with various types of indicators. Simpler indicators are described first. Volume indicators are the simplest of the QIs. Area-level indicators are described next, along with their several possible denominators and the method used to risk-adjust them. Building in complexity, the document describes (1) the calculation of provider-level indicators, in which the denominator is tailored to the indicator and the QI may be affected by the POA data element, and (2) how the software accounts for missing POA data. We describe composite indicators next, and then finish with a description of the methods used to maintain the QI software—specifically the calculations performed to update the reference population and to update denominator data.

Other Helpful Documents

Readers may wish to access additional QI-related documentation. The following are some helpful examples:

QI Software Instructions

SAS: See <http://www.qualityindicators.ahrq.gov/software/SAS.aspx>
WinQI: See <http://www.qualityindicators.ahrq.gov/Software/WinQI.aspx>

QI Technical Specifications

PQI: See http://www.qualityindicators.ahrq.gov/Modules/PQI_TechSpec.aspx
IQI: See http://www.qualityindicators.ahrq.gov/Modules/IQI_TechSpec.aspx
PSI: See http://www.qualityindicators.ahrq.gov/Modules/PSI_TechSpec.aspx
PDI: See http://www.qualityindicators.ahrq.gov/Modules/PDI_TechSpec.aspx

QI Risk-Adjustment Coefficient Tables

PQI: See http://www.qualityindicators.ahrq.gov/modules/pqi_resources.aspx
IQI: See http://www.qualityindicators.ahrq.gov/modules/iqi_resources.aspx
PSI: See http://www.qualityindicators.ahrq.gov/modules/psi_resources.aspx
PDI: See http://www.qualityindicators.ahrq.gov/modules/pdi_resources.aspx

QI Population Documentation File

See <http://www.qualityindicators.ahrq.gov/software/SAS.aspx>
Healthcare Cost and Utilization Project (HCUP) State Inpatient Database (SID)
Documentation (to better understand the source of the reference population): See
<http://www.hcup-us.ahrq.gov/db/state/siddbdocumentation.jsp>

Chapter 1. The User's Dataset

An AHRQ QI software user should prepare the input dataset according to the software instructions.

Table 1.1 Required Data Elements

Data Element	Label	PQI	IQI	PSI	PDI
AGE	Age in years at admission	X	X	X	X
AGEDAY	Age in days (when age <1 year)				X
ASCHED	Admission scheduled vs. unscheduled			X	X
ASOURCE	Admission source (uniform)	X	X	X	X
ATYPE	Admission type			X	X
DISPUNIFORM	Disposition of patient (uniform)		X	X	X
DQTR	Discharge quarter	X	X	X	X
DRG	DRG in effect on discharge date	X	X	X	X
DRGVER	DRG grouper version used on discharge date	X	X	X	X
DSHOSPID	Data source hospital identifier		X	X	X
DX1-DX30	Diagnosis	X	X	X	X
DXPOA1-	Diagnosis present on admission indicator		X	X	X
E_POA1-E_POA10	E code present on admission indicator		X	X	X
ECODE1-	E code		X	X	X
HOSPST	Hospital State postal code		X	X	X
KEY	HCUP record identifier	X	X	X	X
LOS	Length of stay (cleaned)		X	X	X
MDC	MDC in effect on discharge date	X	X	X	X
PAY1	Primary expected payer (uniform)		X	X	X
PAY2	Secondary expected payer (uniform)		X	X	X
POINTOFORIGINU B04	Point of origin for admission or visit, UB-04 standard	X	X	X	X
PR1-PR30	Procedure	X	X	X	X
PRDAY1-	Number of days from admission			X	X
PSTCO	Patient State/county FIPS code	X	X	X	X
PSTCO2	Patient State/county FIPS code, possibly derived from	X	X	X	X
RACE	Race (uniform)	X	X	X	X
SEX	Sex	X	X	X	X
YEAR	Calendar year	X	X	X	X

Abbreviations: DRG, diagnosis-related group; HCUP, Healthcare Cost and Utilization Project; FIPS, Federal Information Processing Standard; MDC, major diagnostic category

Note: The AHRQ QI software deletes discharge records with missing values for SEX.

In preparing a dataset for analysis, data elements and data values shown on the right side of Table 1.2 are constructed from the discharge data elements.

Table 1.2 Data Elements and Data Values to Be Constructed by the User

Discharge Data (e.g., SID)		AHRQ QI	
Data Element	Data Value	Data Element	Data Value
FEMALE	0 – Male 1 – Female	SEX	1 – Male 2 – Female
ATYPE, ASCHED, and AGEDAY	IF ATYPE = Missing AND ASCHED = 1 (Scheduled admission) AND AGEDAY ~= 0	ATYPE	3 – Elective
ECODE1-ECODE10	As reported	DX31-DX40	As reported
E_POA1-E_POA10	As reported	DXPOA31-DXPOA40	As reported

Abbreviations: AHRQ, Agency for Healthcare Research and Quality; QI, Quality Indicator; SID, State Inpatient Databases

Discharge records in the dataset are analyzed as either adult or pediatric data on the basis of age and major diagnostic category (MDC) (Table 1.3). Discharges in MDC 14 (Pregnancy, Childbirth & the Puerperium) are assigned to the adult analysis data regardless of age.

Table 1.3 Analysis Data Inclusion Rule

Analysis Data	Inclusion Rule
Adult	AGE≥18 years or MDC=14
Pediatric	AGE<18 years and MDC≠14

Adult analysis data are used to calculate Prevention Quality Indicators (PQIs), Inpatient Quality Indicators (IQIs), and Patient Safety Indicators (PSIs). Pediatric records are used to calculate Pediatric Quality Indicators (PDIs), Neonatal Quality Indicators (NQIs), and indicators from other modules defined on pediatric discharges (i.e., PQI 09 Low Birth Weight Rate, PSI 17 Birth Trauma Rate – Injury to Neonate).

Chapter 2. Calculating Volume and Count Indicators

This section describes how the software calculates volume and count indicators. Table 2.1 lists the seven volume indicators for inpatient procedures for which there is evidence that a higher volume of procedures conducted by a provider is associated with lower mortality. The volume indicators are measured as counts of hospitalizations in which particular procedures were performed.

Table 2.1 AHRQ QI Volume Indicators

Name
IQI 01 – Esophageal Resection Volume*
IQI 02 – Pancreatic Resection Volume*
IQI 04 – Abdominal Aortic Aneurysm (AAA) Repair Volume*
IQI 05 – Coronary Artery Bypass Graft (CABG) Volume
IQI 06 – Percutaneous Coronary Intervention (PCI) Volume
IQI 07 – Carotid Endarterectomy Volume
PDI 07 – RACHS-1 Pediatric Heart Surgery Volume

*IQI 01, IQI 02, and IQI 04 are intended to be reported with IQI 08, IQI 09, and IQI 11, respectively.

Table 2.2 lists the four count indicators for serious reportable events.

Table 2.2 AHRQ QI Count Indicators

Name
PSI 15 – Retained Surgical Item or Unretrieved Device Fragment Count
PSI 16 – Transfusion Reaction Count
PDI 03 – Retained Surgical Item or Unretrieved Device Fragment Count
PDI 13 – Transfusion Reaction Count

Discharge-Level Indicator Data Element (T)

The terms *numerator* and *denominator* appear throughout the QI documentation. There are no denominators for volume or count indicators. The quantity of interest at the provider level is the magnitude of the number of times the procedure or the event occurs, and that number is not normalized by or divided by any denominator. The technical specifications do, however, use the term *numerator* to define the procedure of interest. Discharge records are flagged for inclusion or exclusion from the numerator of each volume QI on the basis of the data elements, data values, and logic described in the technical specifications for each indicator.

For each discharge record in the dataset, the software calculates a binary flag variable for each volume or count QI. In this document, we denote the discharge-level indicator data element with the letter T (for *top*). The software creates a “T” variable for each QI and the remainder of the variable name corresponds to the identity of the QI (e.g., the variable for IQI 01 is called TPIQ01).

Numerator

Discharges are flagged for inclusion in the numerator of each volume QI according to the specification for the procedure of interest (for volume indicators) or outcome of interest (for count indicators). Discharges flagged for inclusion in the numerator are assigned a value of 1 for T.

Exclusions

The specifications often stipulate that records should be excluded from calculation of a volume indicator if the record is missing an important data element. For volume indicators, excluded records are assigned a value of missing (.) for T.

Observed Value

The observed provider-level value of a volume or count indicator is simply the sum of T (the discharge-level data element) over all records for that provider in the dataset.

Chapter 3. Calculating Area-Level Indicators—Observed Rates

This section describes how the software calculates rates for area-level indicators. *Area-level indicators* contained in the PQI module identify hospital admissions that evidence suggests might have been avoided through access to high-quality community care and resources. The area-level indicators contained in the IQI module identify procedures which may be overused, and the area-level PSI indicators are used to evaluate the frequency of patient safety events geographically. The numerator is a count of admissions for the condition of interest. The denominator is an estimate of the number of persons at risk for such a hospitalization. The denominator is usually a population estimate from a U.S. Census Bureau dataset, but in some cases may be based on the number of discharges. For example, the rate of perforated appendix is calculated with the number of appendectomy discharges as a denominator.

Table 3.1 lists the area-level indicators.

Table 3.1 AHRQ QIs Area-Level Indicators

Name
IQI 26 – Coronary Artery Bypass Graft (CABG) Rate
IQI 27 – Percutaneous Coronary Intervention (PCI) Rate
IQI 28 – Hysterectomy Rate
IQI 29 – Laminectomy or Spinal Fusion Rate
PDI 14 – Asthma Admission Rate
PDI 15 – Diabetes Short-Term Complications Admission Rate
PDI 16 – Gastroenteritis Admission Rate
PDI 17 – Perforated Appendix Admission Rate
PDI 18 – Urinary Tract Infection Admission Rate
PQI 01 – Diabetes Short-Term Complications Admission Rate
PQI 02 – Perforated Appendix Admission Rate
PQI 03 – Diabetes Long-Term Complications Admission Rate
PQI 05 – Chronic Obstructive Pulmonary Disease (COPD) or Asthma in Older Adults Admission Rate
PQI 07 – Hypertension Admission Rate
PQI 08 – Heart Failure Admission Rate
PQI 09 – Low Birth Weight Rate
PQI 10 – Dehydration Admission Rate
PQI 11 – Bacterial Pneumonia Admission Rate
PQI 12 – Urinary Tract Infection Admission Rate
PQI 13 – Angina Without Procedure Admission Rate
PQI 14 – Uncontrolled Diabetes Admission Rate
PQI 15 – Asthma in Younger Adults Admission Rate
PQI 16 – Lower-Extremity Amputation among Patients with Diabetes Rate

The software provides the user with the option of producing output by metropolitan area or by county. The default is to output county-level statistics. The term *metropolitan area* (MA) was adopted by the U.S. Census in 1990 and referred collectively to Metropolitan Statistical Areas (MSAs), Consolidated Metropolitan Statistical Areas, and Primary Metropolitan Statistical Areas. In addition, *area* could refer to (1) Federal Information Processing Standard (FIPS) county, (2) modified FIPS county, (3) 1999 Office of Management and Budget (OMB) MSA, or (4) 2003 OMB MSA. As an aside, Micropolitan Statistical Areas are not used in the QI software.

For information about how the denominators are calculated from census data, see the QI Population Documentation File at <http://www.qualityindicators.ahrq.gov/software/SAS.aspx>.

For diabetes-related area measures, the QI software user has an option of calculating rates in which the denominator is an estimate of the number of persons living in the State who have diabetes. For information on how those *condition-specific denominators* are estimated, see [Chapter 3](#). The diabetes indicators are PQI 01 Diabetes Short-Term Complications Admission Rate, PQI 03 Diabetes Long-Term Complications Admission Rate, PQI 14 Uncontrolled Diabetes Admission Rate, and PQI 16 Lower-Extremity Amputation among Patients with Diabetes Rate. [Chapter 11](#) describes how the diabetes denominators are estimated.

Future versions of the QI software may include other condition-specific denominator options.

Discharge-Level Indicator Data Element (T)

Numerator

The software creates a “T” variables for each QI and the remainder of the variable names identifies the corresponding QI (e.g., the T variable for PQI 01 is called TPPQ01). Discharges are flagged for inclusion in the numerator of each area-level QI according to the specification for the condition of interest. Discharges flagged for inclusion in the numerator are assigned a value of 1 for T.

Exclusions

Generally, discharges may be flagged for exclusion from the numerator of an area-level AHRQ QI for one (or more) of several reasons:

1. The outcome of interest is very difficult to prevent or have an unclear conceptual relationship to access to quality care or community resources.
2. The patient was transferred from another health care facility (to avoid double counting a single encounter).
3. Encounters missing data elements that are required for indicator construction.

Discharge records that meet one or more of the exclusion criteria in the QI technical specification are assigned a value of missing (.) for T.

Observed Rate

The observed rate of an area-level indicator is simply the sum of T (the discharge level indicator data element) over all records for that area of the country divided by the census population estimate for the area (adult population for adult measures and child population for pediatric measures). For condition-specific indicators, if the user requests it, the denominator is the estimated count of persons living in that area of the country who are living with the condition of interest.

Area Rates Stratified by Quarter of the Year

The WinQI software has an option to stratify area-level rates by quarter of the year in which they occurred. When the user selects that option, the rate reported for each quarter is the number of admissions for the condition of interest that occurred during that quarter divided by the census population for the area divided by four. The four quarterly rates sum to the annual rate.

Quarterly rates must be interpreted with caution, given seasonal variation for many PQI and potential decrease in reliability associated with reduced numerator count.

Chapter 4. Risk Adjustment for Area-Level Indicators

In order to make meaningful comparisons of the area-level rate for one area with that of another area, it is helpful to account statistically for differences in demographics between areas. To do so for most QIs, the software calculates a risk-adjusted rate that answers the question: What QI rate would we expect to observe in a particular area of the country if the persons living there shared the same demographic profile of a reference population? In statistical language, the risk adjustment controls for demographic differences via logistic regression.

For area rates, the risk-adjustment models adjust for age-group proportions by sex. That is to say that the models include age (in 5-year groups), sex, and if it is the interaction between age and sex. The PQI module contains an option to incorporate a poverty variable in the risk adjustment model.

When comparing outcomes from different areas, there may be several reasons for differences in risk-adjusted rates. Some of the most important reasons may be related to the availability of quality preventive and outpatient care, and other reasons may contribute as well, but after risk adjustment, the differences should not be attributable to differences in the age and sex profiles in the areas.

AHRQ QI Reference Population

To accomplish risk adjustment, in annual updates of the QI software a reference population is analyzed that consists of HCUP SID data that are available for the year most recently released by AHRQ at the time that the QI software is updated. For example, when version 6.0 of the QI software was updated in January 2016 for the May 2016 software release, SID data were available from 2013 from 44 States, so those records serve as the reference population for AHRQ QI software version 6.0.

For area-level indicators, the reference population plays two important roles:

1. The *reference population rate* for each QI is calculated and included in the software to serve as a comparative standard for areas of the country. One can analyze data to determine which areas have rates that are higher or lower than those of the overall reference population. The reference population rates are published on the AHRQ QI Web site in documents named Benchmark Tables (formerly known as Comparative Data Tables). See the [links](#) in the Overview chapter of this document.
2. The *risk-adjustment models* are re-estimated annually using the most recent reference population dataset. This process is described in [Chapter 11](#) of this document. The models are included in the QI software to allow calculation of risk-adjusted rates. The risk-adjustment model covariates and regression coefficients are published on the AHRQ Web site. See the [links](#) in the Overview chapter of this document.

Chapter 5. Calculating Area-Level Indicators—Expected, Risk-Adjusted, and Smoothed Rates

In addition to observed rates, three other sets of QI rates are calculated for risk-adjusted area-level indicators.

Expected Rate

The *expected rate* for an area-level QI is the rate that would be observed if the amount and quality of outpatient and preventive care available across the reference population were available to individuals living in specific geographic areas. Expected rates are predicted for each area using risk-adjustment model coefficients and covariates that summarize the age and sex distribution of the area's population.

Risk-Adjusted Rate

The AHRQ QI use indirect standardization to calculate risk-adjusted rates. The risk-adjusted rate equals the reference population rate multiplied by the ratio of observed rate divided by expected rate:

$$\text{Risk Adjusted Rate} = \text{Reference Population Rate} \times (\text{Observed Rate} / \text{Expected Rate})$$

Note that for the reference population, the observed rate equals the expected rate equals the reference population rate equals the risk-adjusted rate.

The software estimates the standard error of the risk-adjusted rate for each area using a method recommended by Iezzoni (2013) and described by Hosmer and Lemeshow (1995) that represents the amount of within-provider or area variance due to sampling (i.e., as the number of patients per provider or individuals per area increases, this variance tends to zero). This standard error is used to calculate lower and upper bound 95% confidence intervals around the risk-adjusted rate as risk adjusted rate $\pm 1.96 \times$ risk-adjusted rate standard error (stored in a data element with an L and a U prefix). (See [Chapter 9](#) section titled, Computing the Risk-Adjusted Rate Variance. See also http://qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/Calculating_Confidence_Intervals_for_the_AHRQ_QI.pdf).

Smoothed Rate

Each area's *smoothed rate* is a weighted average of the risk-adjusted rate and the reference population rate calculated from discharges in the reference population; the smoothed rate is calculated with an empirical Bayes shrinkage estimator (1) to result in a rate that will be near that from the input dataset if the area's rate is estimated in a stable fashion with minimal noise or (2) to result in a rate near that of the reference population if the rate from the area is unstable and based on noisy data. Thus, the smoothed rate for an area with stable estimates will be similar to the area's risk-adjusted rate, whereas the smoothed rate for an area with unstable estimates will be similar to the rate calculated in the discharges of the reference population.

The formula for the smoothed rate is as follows:

$$\text{Smoothed Rate} = (\text{Risk adjusted Rate} \times \text{Shrinkage Weight}) + \text{Reference Population Rate} * (1 - \text{Shrinkage Weight}),$$

where

$$\text{Shrinkage Weight} = \frac{\text{Signal Variance}}{\text{Signal Variance} + \text{Noise Variance}}$$

The noise variance is an estimate of variability in the QI outcome within the area of interest (county), and the signal variance is an estimate of variability across all areas.

$$\begin{aligned} \text{Noise Variance } \hat{\sigma}_a^2 &= \left(\frac{\bar{Y}}{n_a E_a} \right)^2 \sum_{i \in A_a} \hat{Y}_i (1 - \hat{Y}_i) \\ \text{Signal Variance } \hat{\tau}^2 &= \frac{\sum_{a=1}^A \frac{1}{(\hat{\tau}^2 + \sigma_a^2)^2} \{ (RAR_a - \overline{RAR})^2 - \hat{\sigma}_a^2 \}}{\sum_{a=1}^A \frac{1}{(\hat{\tau}^2 + \sigma_a^2)^2}} \end{aligned}$$

where A is the number of areas with persons at risk for the measure, \bar{Y} is the observed rate for the reference population; \hat{Y}_i is the person-level predicted probability for area i ; and for area a , A_a is the collection of persons in the population at risk, n_a is the number of persons, E_a is the expected rate, and RAR_a is the risk-adjusted rate. Note that $\hat{\tau}^2$ appears on both sides of the signal variance equation; it is estimated in an iterative fashion (Morris, 1983).

For purposes of confidence interval estimation, the smoothed rate is assumed to follow a gamma distribution $G(\text{shape}, \text{scale})$, to incorporate the influence of the empirical Bayes shrinkage weights on the rates, where

$$\begin{aligned} \text{Shape} &= \frac{(\text{Smoothed Rate})^2}{\text{Posterior Variance}} \\ \text{Scale} &= \frac{\text{Posterior Variance}}{\text{Smoothed Rate}} \end{aligned}$$

$$\text{Posterior Variance} = \text{Signal Variance} - (\text{Shrinkage Weight} * \text{Signal Variance})$$

When there is a fixed comparative rate of interest, it is possible to parameterize the smoothed rate posterior probability on the basis of the gamma distribution and calculate the probability that the smoothed area rate falls below or above the comparative rate that is of interest.

Chapter 6. Overview of Provider-Level Quality Indicators and Present-on-Admission

This section describes how the software calculates provider-level (e.g. hospital-level) indicator rates. *Provider-level indicators* address questions such as: Did the patient experience an adverse quality-related event while in the care of a specific healthcare provider? Or did the patient have an inpatient procedure for which there are questions of overuse, underuse, or misuse?

Adverse-event indicators are for medical conditions and procedures that have been shown to have complication/adverse event rates that vary substantially across institutions and for which evidence suggests that high rates may be associated with deficiencies in the quality of care. Adverse-event indicators usually include only those cases in which a secondary diagnosis code flags a potentially preventable complication. A few indicators are based on procedure codes that imply a potential preventable adverse event.

Mortality indicators are for medical conditions and surgical procedures that have been shown to have mortality rates that vary substantially across institutions and for which evidence suggests that high mortality may be associated with deficiencies in the quality of care.

Utilization indicators track procedures in which there are questions of overuse, underuse, or misuse. The usage of the procedures being examined varies significantly across hospitals and areas, and high or low rates by themselves do not represent poor quality of care; rather the information is intended to inform consumers about local practice patterns.

Provider-level indicators are measured as rates—number of hospitalizations with the outcome (or procedure) of interest divided by the population at risk for the outcome (or procedure). Recall that area-level indicators all use the same denominator for each area—the census-derived estimate of the count of persons who live in the area. Provider-level indicators are more complicated because they have *indicator-specific denominators* to identify only the hospitalizations that were at risk for the outcome of interest.

Recall that area-level indicators all use similar risk-adjustment coefficients: age-group by sex, but the risk-adjustment models for provider-level measures are more complicated. Each risk-adjusted provider-level indicator uses a customized list of regression covariates that are selected when the QI software is updated annually using methods described in [Chapter 11](#).

Present-on-Admission (POA) status is a third factor that makes provider-level indicators more complex than volume- or area-level indicators. Current AHRQ QIs that use POA are listed in [Appendix A](#). Some of the indicators identify adverse conditions that develop as medical complications during the hospitalization of interest. Evidence suggests that high rates may be associated with lower quality of care. Think, for instance, of pressure ulcers, which are measured with PSI 03. However, some of these complications may have been POA, which would not be related to the quality of inpatient care. The AHRQ QI software uses three methods to distinguish between *complications*, which develop during the hospitalization and should be counted in the QI numerator, and *comorbidities*, which are POA and should be excluded from the QI calculation because the patient is not at risk for the event. We summarize those methods in Table 6.1 and describe them in more detail in the following chapters.

Table 6.1 Methods Used by QI Software to Distinguish Complications From Comorbidities

Method	Description	Can the QI User Turn This Off?
1. The POA-Related Exclusion Method (See Chapter 7.)	Some QIs use data elements other than DX_POA to infer that the condition is more likely than not to be POA. Those records are excluded from the population at risk.	No. The WinQI software does not allow modifications to the exclusion criteria. However, the SAS software allows the user to make alterations and notes that the user should document any modifications to the program.
2. DX_POA Data Element (See Chapter 8.)	If the diagnosis is flagged as POA using the DX_POA data element, then the record is excluded from the population of interest.	No. The WinQI software does not allow modifications to the exclusion criteria. However, the SAS software allows the user to make alterations and notes that the user should document any modifications to the program.

Abbreviations: POA, Present on Admission; QI, Quality Indicator

POA Data Element—Background Information

POA was added as a data element to the uniform bill form (UB-04) effective October 1, 2007, and hospitals incurred a payment penalty for not including POA on Medicare records beginning October 1, 2008. Each of the several diagnoses in a discharge record can be flagged as “present at the time the order for inpatient admission occurs” or not (see www.cdc.gov/nchs/data/icd/icd9cm_guidelines_2011.pdf, p. 97). This flag is accomplished with data element DX_POAi, which uses a one-character text code to characterize the POA status of the diagnosis in DXi. Conditions that develop during an outpatient encounter, including treatment in an emergency department, are considered as POA. Most States have adopted POA in the discharge data submitted by hospitals to either the State department of health or the State hospital association.

Table 6.2 lists the possible character values of the POA data elements (Y, N, U, W, E, or missing) along with corresponding numeric values (0 or 1) used in the AHRQ QI software. Additional information about the coding guidelines for POA can be found at www.cdc.gov/nchs/data/icd/icd9cm_guidelines_2011.pdf. Again, current AHRQ QIs that use POA are listed in [Appendix A](#).

Table 6.2 Values for the Present-on-Admission Data Element

ICD-9-CM Guidelines	Description	AHRQ QI POA Data Element	Description
Y - Yes	Diagnosis is present at the time of inpatient admission	1	Diagnosis present at admission
N – No	Diagnosis is not present at the time of inpatient admission	0	Diagnosis not present at admission
U – Unknown	Documentation is insufficient to determine whether condition is present on admission	0	Diagnosis not present at admission
W – Clinically undetermined	Provider is unable to clinically determine whether condition is present on admission	1	Diagnosis present at admission
E – Unreported/not used; also includes UB-04 values previously coded as 1	Reported as exempt from reporting on a nonexempt diagnosis	0	Diagnosis not present at admission
X – End of POA indicators	Denotes the end of the POA indicators (terminated 1/2011)	0	Diagnosis not present at admission

Abbreviations: AHRQ, Agency for Healthcare Research and Quality; POA, Present on Admission; QI, Quality Indicator

Source (<https://www.cms.gov/medicare/medicare-fee-for-service-payment/hospitalacqcond/coding.html>); Healthcare Cost and Utilization Project (HCUP), Agency for Healthcare Research and Quality. Central Distributor SID: Description of Data Elements. http://www.hcup-us.ahrq.gov/db/vars/siddistnote.jsp?var=e_poan. Accessed November 10, 2015.

An individual discharge record might include 20 or more diagnoses. For purposes of the AHRQ QI, the principal diagnosis is always assumed to be POA by definition, regardless of the coding of the POA data element in the principal field. Secondary diagnosis codes first are checked to see whether the diagnosis is exempt from reporting POA. If the secondary diagnosis is exempt, it is considered POA. If the secondary diagnosis is not exempt, then it considered POA if the POA data element is coded with a Y or W. Secondary diagnosis codes are considered not POA if the POA data element is coded with an N, a U, a blank, an E, a 1, or an X. The AHRQ QI software assumes that POA information is present and accurately coded.

Chapter 7. Calculating Provider-Level Observed Rates

Discharge-Level Indicator Data Element (T)

For each discharge record in the dataset, the software calculates a binary flag variable for each provider-level QI. In this document, we denote the discharge-level indicator data element with the letter T (for *top*). The software creates a “T” variables for each QI and the remainder of the variable name identifies the corresponding QI (e.g., the T variable for IQI 09 is called TPIQ09).

Each provider-level observed QI rate consists of a conceptually simple fraction in which the denominator is the count of discharge records at risk and the numerator is the count of records with the outcome of interest. This fraction is calculated using a single discharge-level indicator data element, T, described in earlier chapters for volume- and area-level indicators. In those earlier chapters, the T variable took on the value 1 if the discharge record met the definition for the numerator that is spelled out in the technical specifications. For volume- and area-level indicators, it does not matter whether the T variable takes the value 0 or missing (.) for other records, because the numerator is simply the count of records where T=1.

Provider-Level Denominator

Discharges are flagged for inclusion in the denominator of each AHRQ QI according to the specification for the population at risk. Discharges flagged for inclusion in the denominator are assigned a value of 0 for T unless the discharge also experienced the outcome of interest, in which case the value of 1 is assigned. Discharges that experienced the outcome of interest are in the population at risk by definition.

Denominator Exclusions

Generally, discharges may be flagged for exclusion from the denominator of an AHRQ QI for one (or more) of several reasons:

1. The outcome of interest has been coded as present on admission.
2. The outcome of interest is very difficult to prevent and therefore not an indication of substandard care.
3. The exclusion identifies populations who are at very low risk for the adverse event and who are excluded to keep from diluting the QI denominator.
4. Some exclusion criteria are included for the purpose of enhancing face validity with clinicians (e.g., exclude patients from being at risk of a pressure ulcer [PSI 03] if they have not been hospitalized for at least 3 days).
5. Some exclusion criteria are an inherent part of the QI definition.

Discharge records that meet one or more of the denominator exclusion criteria in the QI technical specification are assigned a value of missing (.) for T.

Three Values of T (Discharge Level Indicator Data Element)

To summarize:

- A 1 in the T variable means that the record was in the population at risk, experienced the outcome of interest, and was not excluded for any reason.
- A 0 in the T variable means the record was in the population at risk, but did not experience the outcome of interest, and was not excluded for any reason.
- A missing (.) value for the T variable means that the record was not in the population of interest, either because it did not meet the denominator definition or because it met one or more of the exclusion criteria.

The Observed Rate

For provider-level indicators, the observed rate is simply the arithmetic mean of the T variable over all of the provider's discharge records.

Discharge-Level POA Exclusion Data Element (Q)

Consideration of present on admission (POA) status should improve the accuracy of QI rate calculation because pre-existing comorbidities can be distinguished from complications that develop during the hospital stay of interest. Records with outcomes that were POA no longer will appear erroneously in the numerator, denominator, or observed rate, and the risk-adjustment models no longer will erroneously treat complications as comorbidities, thus yielding improvement in the comparative expected, risk-adjusted, and smoothed rates above and beyond that in the numerator, denominator, and observed rates.

The degree of improvement attained when accounting for POA will vary (1) depending on the number of records in which the outcomes were POA and (2) with the accuracy of POA coding. This document does not address the topic of POA accuracy. The QI software treats eligible values in the DX_POA data elements as if they were completely accurate. Values that are ineligible or missing are treated as if they were coded as "not present on admission." In other words, missing or ineligible values lead to a condition being treated as a complication.

When accounting for POA, the QI software codes the discharge-level indicator data element, T, in the same manner described above, using technical specifications to define which records are included in the denominator or the numerator and which should be excluded for one or more reasons.

A second, POA-related, binary flag is calculated also. The discharge-level POA exclusion data element is abbreviated with the letter Q.¹ Put simply, Q records whether the outcome of interest

¹ The letter *P* was not available, having been used already for the notion of population at risk. In this

was POA or not. The outcome of interest is considered POA (Q is assigned 1) if any of the diagnosis codes that define the outcome of interest are coded as POA. Otherwise a value of 0 is assigned to Q. For every record that includes POA data in the SID DX_POA data elements, Q will have a value of 0 or 1 and will not be missing (.).

Observed Rate

Before calculating the observed rate, Q is used to correct the value of T if the condition of interest was POA. If the value of Q is 1 (outcome was POA) then the record is removed from the population at risk by setting T to missing (.). The observed rate is simply the arithmetic mean of the T variable after this correction.

document, the variables are denoted simply as T and Q, but each discharge record has a binary T variable and a binary Q variable for each QI, so the variables have longer names to clarify which QI they describe.

Chapter 8. Risk Adjustment for Provider-Level Indicators

This chapter describes risk-adjustment for provider-level QIs. Provider-level indicators are risk adjusted in a manner similar to that described in Chapters [4](#) and [5](#) for area-level indicators. One important difference is that the list of covariates for provider-level indicators differs from indicator to indicator more than those for the area-level indicators. In the next section, we describe the types of data elements that are considered as potential risk adjusters.

Where possible, the logistic regression models use a generalized estimating equations (GEEs) approach to account for correlation at the provider (i.e., hospital) level. When GEE models do not converge during the annual AHRQ QI software update, then multivariable logistic regression models are used that do not account for within-provider correlation. See [Chapter 11](#) for more details.

Risk-Adjustment Covariates

Each risk-adjusted QI (listed in [Appendix A](#)) has a set of covariates that have been identified as useful covariates in a logistic regression risk-adjustment model. [Chapter 11](#) describes the variable selection process.

For the PSIs, covariates include sex, age, modified diagnosis-related group (MDRG), major diagnostic category (MDC), a set of HCUP comorbidities, and transfer status.

For the IQIs, include sex, age, all patient refined diagnosis related groups (APR-DRGs), and risk-of-mortality subclass (minor, moderate, major, extreme) that are used as covariates in the risk-adjustment model along with transfer status.

For the PDIs, include birth weight, age in days, age in years, modified diagnosis-related group (MDRG), at least 1 of 46 clinical classification software (CCS) comorbidities, and some indicator-specific risk categories that are used as covariates in the risk-adjustment model.

Risk Adjustment Parameters CSV File

Each risk-adjusted provider-level indicator has its risk adjustment parameter estimates stored in a comma separated values (.csv) file that accompanies the QI software.

Expected Rate

Using the risk adjustment parameters, each eligible discharge (i.e., one that is included in the denominator of the indicator) is scored for its expected (or predicted) rate using PROC SCORE. This output score is simply the sum across all covariates in the risk-adjustment model of the scalar multiplication of the presence or absence of a covariate (1 or 0) times the value of the coefficient from the risk-adjustment model for that covariate. This score is the logit of the

predicted value (denoted MHAT in the software). The predicted probability for the discharge is computed as follows:

$$EHAT = \exp(MHAT) / [1 + \exp(MHAT)]$$

The discharge-level predicted probabilities are used to compute an expected rate for the indicator by:

$$\text{Expected Rate} = \frac{\text{Sum of the predicted rates for each discharge}}{\text{Count of discharges in the population at risk}}$$

The Risk-Adjusted Rate

The AHRQ QIs use indirect standardization to calculate the risk-adjusted rate:

$$\text{Risk Adjusted Rate} = \text{Reference Population} \times (\text{Observed Rate} / \text{Expected Rate})$$

The software estimates the standard error of the risk-adjusted rate for each provider or area using a method recommended by Iezzoni (2013) and described by Hosmer and Lemeshow (1995) that represents the amount of within-provider or area variance due to sampling (i.e., as the number of patients per provider or individuals per area increases, this variance tends to zero). This standard error is used to calculate lower and upper bound 95% confidence intervals around the risk-adjusted rate as risk-adjusted rate $\pm 1.96 \times$ risk adjusted rate standard error (stored in a data element with an L and a U prefix). (See the note below titled Computing the Risk-Adjusted Rate Variance. See also

http://qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/Calculating_Confidence_Intervals_for_the_AHRQ_QI.pdf).

Smoothed Rate

Each provider's smoothed rate is a weighted average of the risk-adjusted rate and the reference population rate calculated from discharges in the reference population; the smoothed rate is calculated with an empirical Bayes shrinkage estimator (1) to result in a rate that will be near that calculated from the input dataset if the provider's rate is estimated in a stable fashion with minimal noise, or (2) to result in a rate near that of the reference population if the rate from the provider is unstable and based on noisy data. Thus, the smoothed rate for a provider with stable estimates will be similar to the provider's risk adjusted rate, whereas the smoothed rate for a provider with unstable estimates will be more similar to the rate calculated in the discharges of the reference population.

The formula for the smoothed rate is as follows:

$$\begin{aligned} \text{Smoothed Rate} = & (\text{Risk Adjusted Rate} \times \text{Shrinkage Weight}) \\ & + \text{Reference Population Rate} \times (1 - \text{Shrinkage Weight}) \end{aligned}$$

where

$$\text{Shrinkage Weight} = \frac{\text{Signal Variance}}{\text{Signal Variance} + \text{Noise Variance}}$$

The noise variance is calculated for each hospital based on the user's data. The signal variance is a parameter calculated from the reference population. Beginning in version 4.3, there are two signal variance estimates: one using POA and one ignoring POA data.

$$\begin{aligned} \text{Noise Variance } \hat{\sigma}_h^2 &= \left(\frac{\bar{Y}}{n_h E_h} \right)^2 \sum_{i \in A_h} \hat{Y}_i (1 - \hat{Y}_i) \\ \text{Signal Variance } \hat{\tau}^2 &= \frac{\sum_{h=1}^H \frac{1}{(\hat{\tau}^2 + \sigma_h^2)^2} \{ (RAR_h - \overline{RAR})^2 - \hat{\sigma}_h^2 \}}{\sum_{h=1}^H \frac{1}{(\hat{\tau}^2 + \sigma_h^2)^2}} \end{aligned}$$

where H is the number of hospitals with patients at risk for the QI, \bar{Y} is the observed rate for all discharges in the reference population; \hat{Y}_i is the patient-level predicted probability; and for hospital h , A_h is the collection of patients, n_h is the number of patients, E_h is the expected rate, and RAR_h is the risk-adjusted rate. Note that $\hat{\tau}^2$ appears on both sides of the signal variance equation; it is estimated in an iterative fashion (Morris, 1983).

For purposes of confidence interval estimation, the smoothed rate is assumed to follow a gamma distribution $G(\text{shape}, \text{scale})$ to incorporate the effect of the empirical Bayes shrinkage weight, where

$$\begin{aligned} \text{Shape} &= \frac{(\text{Smoothed Rate})^2}{\text{Posterior Variance}} \\ \text{Scale} &= \frac{\text{Posterior Variance}}{\text{Smoothed Rate}} \end{aligned}$$

$$\text{Posterior Variance} = \text{Signal Variance} - (\text{Shrinkage Weight} * \text{Signal Variance})$$

When there is a fixed comparative rate of interest, it is possible to parameterize the smoothed rate posterior probability of the gamma distribution and calculate the probability that the smoothed area rate falls below or above the comparative rate that is of interest.

Computing the Risk-Adjusted Rate Variance

Let

Y_i be the observed (0, 1) outcome for patient i

E_i be the expected (predicted) rate

n_h be the number of discharges at hospital h

α be the reference population rate (average outcome in the entire sample)

We define the observed rate at hospital h as

$$O_h = \frac{1}{n_h} \sum_{i=1}^{n_h} Y_i$$

the expected rate at hospital h as

$$E_h = \frac{1}{n_h} \sum_{i=1}^{n_h} E_i$$

and the risk-adjusted rate

$$RAR_h = \alpha \times \frac{O_h}{E_h}$$

Using a Taylor expansion for the formula for the variance of the ratio of two stochastic variables R, S (delta method),

$$Var\left(\frac{R}{S}\right) \cong \frac{E[R]^2}{E[S]^2} \left(\frac{Var(R)}{E[R]^2} - 2 \frac{Cov(R, S)}{E[R]E[S]} + \frac{Var(S)}{E[S]^2} \right)$$

We compute the variance on the risk-adjusted rate:

$$Var(RAR_h) \cong \alpha^2 \frac{E[O_h]^2}{E_h^2} \left(\frac{Var(O_h)}{E[O_h]^2} - 2 \frac{Cov(O_h, E_h)}{E[O_h]E_h} + \frac{Var(E_h)}{E_h^2} \right)$$

It is common practice in these calculations to neglect the variance of the predictor E_h (Hosmer & Lemeshow, 1995) and to consider a normal distribution for the risk-adjusted rate (only true in the limit $n_h \rightarrow \infty$). In this case, the above formula simplifies to

$$Var(RAR_h) \cong \alpha^2 \frac{Var(O_h)}{E_h^2}$$

and the 95% confidence intervals are calculated assuming normality. However, arguments to support using nonapproximate equations (see Luft & Brown, 1993, for an example) for the RAR confidence intervals (in particular when n_h is small) may be considered in future releases of the AHRQ QI software.

Computing the Smoothed Rate Variance

The detailed formula for calculating the probability interval around the smoothed rate is described in [Chapter 9](#) on composite measures. Calculation of the smoothed rate is a step in the process of computing the composite measures. However, the basic formula is as follows:

$$\text{Smoothed Rate} = (\text{Risk Adjusted Rate} \times \text{Shrinkage Weight})$$

$$+ \text{Reference Population Rate} * (1 - \text{Shrinkage Weight})$$

$$\text{Shrinkage Weight} = \frac{\text{Signal Variance}}{\text{Signal Variance} + \text{Noise Variance}}$$

$$\text{Posterior Variance} = \text{Signal Variance} - (\text{Shrinkage Weight} * \text{Signal Variance})$$

The *smoothed rate* follows a gamma distribution $G(\text{shape}, \text{scale})$, where

$$\text{Shape} = \frac{(\text{Smoothed Rate})^2}{\text{Posterior Variance}}$$

$$\text{Scale} = \frac{\text{Posterior Variance}}{\text{Smoothed Rate}}$$

When there is a fixed comparative rate of interest, it is possible to parameterize the posterior probability on the basis of the gamma distribution and calculate the probability that the smoothed area rate falls below or above the comparative rate that is of interest.

Chapter 9. Estimating Composite Measures

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 as described below. The second composite is a weighted average of the component indicators, where the weights are selected on the basis of the intended use of the composite measure. These weights might be determined empirically or based on nonempirical considerations.

Composite Value

The basic steps for computing the composite are as follows:

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

The AHRQ QI risk-adjusted rate and confidence interval are computed as described above.

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

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

Step 3. Compute the reliability-adjusted ratio.

The reliability-adjusted ratio 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).

$$\begin{aligned} & \text{Reliability adjusted ratio (risk – adjusted ratio} \times \text{weight)} \\ & + \text{reference population ratio} \times (1 - \text{weight}). \end{aligned}$$

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. The AHRQ QI software user has the ability to modify these weights in the software, either in the SAS code or in the WinQI user interface. Options for weights include the following:

- *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.
- *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 also might use weights that reflect the amount of excess mortality or complications associated with the adverse event or the amount of confidence that 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 degree 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 an 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 for each composite are based on a principal components factor analysis of the reliability-adjusted ratios.
- *Harm weight.* Harm weighting is based on an analysis that assigns each component indicator a weight that reflects the contribution of that indicator to excess harmful outcomes that occur in the population that experience the component events. Component indicators that both are common and lead to significant excess mortality and morbidity will have the highest weights, whereas those that are less common or have lower mortality and morbidity associated with them will have lower weights. See “Calculating harm weights for the PSI-90 composite” for additional information.

Note: The IQI composites (IQI 90 and 91) use denominator weights, the PDI composites (PSI 90 and PDI 19) use numerator weights, and the PSI composite (PSI 90) uses harm weights.

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.

$$\begin{aligned} \text{Composite} = & (\text{indicator}_1 \text{RAR} \times \text{weight}_1) \\ & + (\text{indicator}_2 \text{RAR} \times \text{weight}_2) + \cdots + (\text{indicator}_N \text{RAR} \times \text{weight}_N) \end{aligned}$$

Calculating harm weights for the PSI-90 Composite

The PSI composite combines smoothed (empirical Bayes shrunken) standardized morbidity ratios (observed/expected ratios) from selected AHRQ Patient Safety Indicators (PSIs) to provide a composite that gives an overview of hospital level quality as it relates to a set of hospital-related events that are associated with harmful outcomes for patients. In past versions of the AHRQ QI software PSI-90 (v5.0 and earlier) the weight that each component received was proportional to the volume of the events in the component indicator observed in the HCUP reference population (i.e. numerator weighting). The re-weighting of PSI-90 was undertaken to improve the validity and reliability of the composite by refining the component indicators that are included in the composite and aligning the weights with the burden of harm (risk of harmful outcomes) that each component contributes in a reference population. In other words, the new weights account for both the magnitude of harm associated with a patient safety event as well as the volume (number of cases) of the event, whereas in past iterations only the volume was used for weighting.

The new weights are defined and calculated as follows:

Each component PSI indicator, q , which is part of PSI 90 receives a weight defined by:

$$weight_q = \frac{volume_q \sum_{h=1}^H harm_{qh} disutility_{qh}}{\sum_{q=1}^Q volume_q \sum_{h=1}^H harm_{qh} disutility_{qh}}$$

Where:

Q is the total number of component quality indicators, q , in PSI 90.

H is the total number of outcome types (harms), h , related to each component indicator.

volume is the numerator count, or the number of total QI events within the component indicator in the reference population.

harm is the excess risk (risk difference) of each type of outcome (i.e. harm) within each component indicator estimated from a model comparing people with PSI events to those without PSI events in an “at risk” cohort.

disutility is the complement of a utility weight (1-utility_wt) assigned to each excess occurrence of each type of outcome within each component indicator.

For each component indicator in the PSI 90 composite, two sets of values need to be computed or estimated. The first is the excess risk of the outcomes (risk difference) that may occur as a consequence of the patient safety event associated with the indicator. The second is the set of numerator weights. Although estimates of disutility are required to incorporate disparate types of harms, the values of disutility are treated as not varying.

Estimating Excess Harms

The estimates of excess harms that go into the harm weighting aim to answer the question, how much more likely is a particular harmful outcome in a population of patients who experience a PSI event than in a population of patients who were at risk for the event, but did not experience the event. In other words, what is the risk difference between PSI events and non-events in an at-risk population? These models require the use of longitudinal data that contain information about morbidity and mortality following a PSI event.

For version 6.0 of the software, excess harms were modeled using CMS Inpatient and Outpatient Medicare Fee-For-Service data in the 100% standard analytical files (SAF). A separate cohort sample was defined for each component indicator based on the sample of 2012 patient records who were “at risk” (i.e. in the denominator) for the component QI indicator. Index events were identified as patient discharges in 2012 with an eligible QI PSI component event. The comparison group was composed of at risk patients (as defined by the component PSI specification) who did not experience the PSI event. The 2013 data were used solely to provide follow-up information about harms. The follow-up period was one year from the discharge date of the index hospitalization. For each component indicator, the independent variable was the T flag for the component PSI event. Separate models were fit for each harm outcome. Outcomes varied among the component PSIs. Example outcomes included all-cause 30-day and 180-day mortality, hospital readmissions, condition-specific complications, and total length of hospital stay (potentially including the postoperative period during the index admission plus all qualifying readmissions within the ascertainment window). The selection of outcomes relied on the underlying conceptual model for the component indicator, the available data elements in the CMS data, and the availability of a meaningful utility weight.

Confounding may arise if factors associated with the probability of experiencing a QI event are also related to the probability of experiencing a consequence (outcome) from the QI event. To account for potential confounding in these analyses, for each component indicator, we used a propensity score weighting approach. The propensity score (PS) was the predicted value (i.e. expected value) from the QI’s risk adjustment model, which accounted for age and sex as well as pre-existing complications and comorbidities. We used a version of propensity weighting suitable for estimating the average treatment effect on the treated (ATT). In other words, we estimated the effect of the safety event on harms among patients who suffer the safety event. Patient stays with the safety event (QI=1) received a weight of 1 and at-risk patient stays without a safety event (QI=0) received a weight of $PS/(1-PS)$.

Another potential source of confounding may arise from patients who experience multiple PSI events that share common outcomes (e.g. mortality). In this scenario it is necessary to estimate independent associations between PSI events and outcomes. When multiple component PSIs are related to the same outcome, we included the other component PSIs in the model as covariates for the excess harm effect we were estimating. For example, if we are estimating the excess risk of renal failure in PSI13, we would use propensity weights appropriate for PSI13 and would also include PSI10 as an indicator covariate in the model.

Harm utility values

To combine disparate harms into a single overall weight, we applied disutility values that scale the relative utility of health states from a patient perspective. Utilities were anchored at zero for

mortality and one for no harmful health outcome. When available, intermediate utility values were drawn from studies that examine patient preference for various health states (e.g. standard gamble studies). When literature-based utility values were not available for patient preference, we used an expert panel of clinicians (physicians and nurses) to rank a list of health states that they have seen in their patients. We applied a regression process to interpolate utility values based on the consensus ranking of the health states. Disutility was calculated as the complement of utility (i.e. 1-utility).

Final PSI-90 weights

The final PSI-90 weights were computed using the excess harm and disutility values derived from the procedures above and combined with information about the volume of the PSI-90 components (T-flags) in the 2013 reference population. The v6.0 AHRQ QI software contains two sets of weights for PSI-90. The first is optional and based on 11 component PSI indicators (PSI 03, and PSI 06 – PSI 15). The second set of weights is the default configuration and these weights have PSI 07 set to zero and the remaining component weights re-scaled to sum to 1.0.

Table 9.1 Weights of PSI-90 component indicators

Table 9.1 Contributions of harms and volume to component PSI weights in the v6.0 PSI-90 composite.											
	PSI 03	PSI 06	PSI 07	PSI 08	PSI 09	PSI 10	PSI 11	PSI 12	PSI 13	PSI 14	PSI 15
Harm Summary (Sum of Excess Harms * Disutilities)	0.258	0.346	0.151	0.252	0.138	0.440	0.215	0.189	0.477	0.133	0.236
Harm Weight (%)	8.5%	11.4%	5.0%	8.3%	4.6%	14.5%	7.1%	6.2%	15.7%	4.4%	7.8%
Volume (Numerator Count)	2957	5898	2305	162	22825	2342	21015	20438	10612	1402	724
Volume Weight (%)	3.3%	6.5%	2.5%	0.2%	25.2%	2.6%	23.2%	22.5%	11.7%	1.5%	0.8%
Harm Score (Harm Summary * Volume)	762.17	2042.54	347.46	40.84	3152.09	1031.14	4519.28	3866.05	5062.24	186.69	170.89
Final Weight (%)	3.6%	9.6%	1.6%	0.2%	14.9%	4.9%	21.3%	18.3%	23.9%	0.9%	0.8%
Final Weight without PSI-07 (%)	3.7%	9.8%	0.0%	0.2%	15.1%	4.9%	21.7%	18.5%	24.3%	0.9%	0.8%
*Source: CMS Inpatient and Outpatient 100% Standard Analytical Files 2012 and 2013 for excess harm models. AHRQ HCUP 2013 QI POA Reference Population for volume											

Composite Variance

The probability interval of the composite measure is based on its standard error, which is the square root of the variance. The variance is computed based on the signal variance-covariance matrix and the reliability weights.

Let \mathbf{M} be a $1 \times K$ vector of observed quality measures (for a given hospital, suppress hospital subscript for convenience), noisy measures of the true underlying $1 \times K$ quality vector $\boldsymbol{\mu}$, such that:

$$\mathbf{M} = \boldsymbol{\mu} + \boldsymbol{\epsilon} \quad (11.1)$$

where $\boldsymbol{\epsilon}$ is a $1 \times K$ noise vector with zero mean and $K \times K$ variance-covariance matrix $\text{Var}(\boldsymbol{\epsilon}) = \boldsymbol{\Omega}_{\epsilon}$. Let the $K \times K$ signal variance-covariance be $\text{Var}(\boldsymbol{\mu}) = \boldsymbol{\Omega}_{\mu}$.

Let $\hat{\boldsymbol{\mu}}$ be a $1 \times K$ vector indicating the posterior (filtered) estimate of $\boldsymbol{\mu}$, such that:

$$\hat{\boldsymbol{\mu}} = \boldsymbol{\mu} + \mathbf{v} \quad (11.2)$$

where \mathbf{v} is a $1 \times K$ vector with zero mean and $K \times K$ variance-covariance matrix $\text{Var}(\mathbf{v})$ representing the prediction error of the posterior estimates.

The goal is to estimate the variance for any weighted average of the posterior estimates. For a given $1 \times K$ weighting vector \mathbf{w} , this is given by:

$$\text{Var}(\mathbf{w}\mathbf{v}) = \mathbf{w}' \text{Var}(\mathbf{v}) \mathbf{w},$$

where \mathbf{w}' indicates the transpose of \mathbf{w} .

Thus, we need an estimate of $\text{Var}(\mathbf{v})$. We simplify the calculation by assuming that the filtered estimates are formed in isolation for each measure (univariate) and that 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).

Forming each measure in isolation, using superscripts $k = 1, \dots, K$ to indicate the measure, we have:

$$\hat{\mathbf{u}}^k = \mathbf{M}^k \hat{\boldsymbol{\beta}}^k = \mathbf{M}^k (\boldsymbol{\Omega}_{\mu}^{kk} + \boldsymbol{\Omega}_{\epsilon}^{kk})^{-1} \boldsymbol{\Omega}_{\mu}^{kk}$$

$$\text{Var}(\mathbf{v}^k) = \boldsymbol{\Omega}_{\mu}^{kk} (1 - \hat{\boldsymbol{\beta}}^k) = \boldsymbol{\Omega}_{\mu}^{kk} - \boldsymbol{\Omega}_{\mu}^{kk} (\boldsymbol{\Omega}_{\mu}^{kk} + \boldsymbol{\Omega}_{\epsilon}^{kk})^{-1} \boldsymbol{\Omega}_{\mu}^{kk},$$

where

$$\hat{\boldsymbol{\beta}}^k = (\boldsymbol{\Omega}_{\mu}^{kk} + \boldsymbol{\Omega}_{\epsilon}^{kk})^{-1} \boldsymbol{\Omega}_{\mu}^{kk}$$

is the signal ratio of measure k , the reliability of the measure, and is the r -squared that measures how much of the variation in the true measure can be explained with the filtered measure. Note that in this simplified case the filtered estimate is a univariate shrinkage estimator. For the nondiagonal elements of the covariance matrix (for $j \neq k$),

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

assuming independent estimation error in the two measures, one gets the following simplified expression (see supplemental notes below for the derivation):

$$Cov(\mathbf{v}^j, \mathbf{v}^k) = \boldsymbol{\Omega}_{\mu}^{jk}[(1 - \hat{\beta}^j)(1 - \hat{\beta}^k)]$$

Note that this is just the signal covariance times 1 minus the signal ratio for each of the measures. Thus, if the signal ratio is 0 for each measure, the covariance in the estimates is simply the signal covariance. As either measure gets a stronger signal ratio (becomes more precise), the covariance in the estimates shrinks to 0.

Also note that if one measure is missing, then the signal ratio is simply set to 0. The filtered estimate is shrunk all the way back to the (conditional) mean, and the variance and covariance are as defined above.

The standard error on the composite is the square root of the variance, which is then used to compute the 95% probability interval.

The composite value follows a gamma distribution $G(shape, scale)$, where

$$Shape = \frac{(Composite\ Value)^2}{Posterior\ Variance}$$

$$Scale = \frac{Posterior\ Variance}{Composite\ Value}$$

A 95% probability interval can be calculated using the inverse CDF of the gamma distribution as

$$Lowerbound = inv_cdf_gamma(0.025, shape, scale)$$

$$Lowerbound = inv_cdf_gamma(0.975, shape, scale)$$

Supplemental Notes:

To derive formula (11.6), we substitute

$$\hat{\mu} = \mathbf{M}\hat{\beta} = (\mu + \epsilon)\hat{\beta}$$

into (11.5) and obtain (for $j \neq k$)

$$\begin{aligned} Cov(\mathbf{v}^j, \mathbf{v}^k) &= E[(\mu^j - (\mu^j + \epsilon^j)\hat{\beta}^j)(\mu^k - (\mu^k + \epsilon^k)\hat{\beta}^k)] \\ &= E[(\mu^j(1 - \hat{\beta}^j) - \epsilon^j\hat{\beta}^j)(\mu^k(1 - \hat{\beta}^k) - \epsilon^k\hat{\beta}^k)] \\ &= E[\mu^j\mu^k(1 - \hat{\beta}^j)(1 - \hat{\beta}^k) + \mu^k\epsilon^j(1 - \hat{\beta}^k)\hat{\beta}^j + \mu^j\epsilon^k(1 - \hat{\beta}^j)\hat{\beta}^k + \epsilon^j\epsilon^k\hat{\beta}^j\hat{\beta}^k] \\ &= E[\mu^j\mu^k](1 - \hat{\beta}^j)(1 - \hat{\beta}^k) + E[\mu^k\epsilon^j](1 - \hat{\beta}^k)\hat{\beta}^j + E[\mu^j\epsilon^k](1 - \hat{\beta}^j)\hat{\beta}^k + E[\epsilon^j\epsilon^k]\hat{\beta}^j\hat{\beta}^k. \end{aligned}$$

Assuming $E[\mu^j\mu^k] = E[\epsilon^j\epsilon^k] = E[\epsilon^j\epsilon^k] = 0$ and $E[\mu] = 0$, we have

$$\begin{aligned} Cov(\mathbf{v}^j, \mathbf{v}^k) &= E[\mu^j\mu^k](1 - \hat{\beta}^j)(1 - \hat{\beta}^k) \\ &= Cov(\mu^j, \mu^k)(1 - \hat{\beta}^j)(1 - \hat{\beta}^k) - E[\mu^j]E[\mu^k](1 - \hat{\beta}^j)(1 - \hat{\beta}^k) \\ &= Cov(\mu^j, \mu^k)(1 - \hat{\beta}^j)(1 - \hat{\beta}^k). \end{aligned}$$

QED.

Chapter 10. Specifications for AHRQ QI Reference Population

In order to maintain the scientific acceptability of the AHRQ QI, the indicators are updated annually to reflect the Uniform Bill (UB-04) coding updates effective each year on July 1, and the International Classification of Diseases-Ninth Revision-Clinical Modification (ICD-9-CM) and Medicare severity diagnosis-related group (MS-DRG) coding updates effective each fiscal year on October 1 of the prior year. In addition, the annual updates include new census data on the population of counties and new HCUP data for the reference population and risk-adjustment covariate coefficients. In this chapter, we describe the methods used to update the QI reference population and the associated risk-adjustment covariate coefficients.

For the version 6.0 release (May 2016), the AHRQ QI program used the HCUP SID for 2013 to compute reference population data. HCUP is a family of health care databases and related software tools and products developed through a Federal-State-Industry partnership and sponsored by AHRQ. HCUP databases bring together the data collection efforts of State data organizations, hospital associations, private data organizations, and the Federal government to create a national information resource of encounter-level health care data. HCUP includes the largest collection of longitudinal hospital care data in the United States, with all-payer, encounter-level information beginning in 1988. These databases enable research on a broad range of health policy issues, including cost and quality of health services, medical practice patterns, access to health care programs, and outcomes of treatments at the national, State, and local market levels. The HCUP SID encompass about 97 percent of all annual inpatient discharges in the United States.

The reference population file was limited to community hospitals and also excludes rehabilitation and long-term acute care hospitals. Information on the type of hospital was obtained by the American Hospital Association (AHA) Annual Survey of Hospitals. AHA defines community hospitals as “all non-Federal, short-term, general, and other specialty hospitals, excluding hospital units of institutions.” Included among community hospitals are specialty hospitals such as obstetrics-gynecology, ear-nose-throat, orthopedic, and pediatric institutions. Also included are public hospitals and academic medical centers.

The 2013 HCUP SID includes information on all inpatient discharges from hospitals in 40 participating States. In 2013, 34 of the SID include indicators of the diagnoses being present on admission (POA) and included the PRDAY data element. Edit checks on POA were developed during an HCUP task that examined POA coding in the 2011 SID at hospitals that were required to report POA to CMS. The edits identify general patterns of suspect reporting of POA. The edits do not evaluate whether a valid POA value (e.g., Y or N) is appropriate for the specific diagnosis. There are three hospital-level edit checks:

1. Indication that a hospital has POA reported as Y on all diagnoses on all discharges
2. Indication that a hospital has POA reported as missing on all non-Medicare discharges
3. Indication that a hospital reported POA as missing on all nonexempt diagnoses for 15 percent or more of discharges. The cut-point of 15 percent was determined by 2 times the standard deviation plus the mean of the percentage for hospitals required to report POA to CMS.

There are several important steps in the annual update process upstream from risk adjustment and rate estimation. Changes may be made to QI technical specifications for one reason or another. Those must be implemented in the software. ICD-9 code sets may be modified. Those changes need to be updated in the software as well. The software is designed to be backward compatible, applying the appropriate sets of codes to older datasets. This work is accomplished before risk-adjustment models are calculated.

Estimating risk-adjustment models and calculating QI rates in the reference population involves running the QI software on the reference population dataset. Note that calculations for area-level indicators are run using the full reference population and calculations for provider-level indicators are run using the subset of the reference population with good POA information as outlined above.

Chapter 11. Estimating Risk Adjustment Models

Construction of Candidate Covariates for Risk Adjustment

For the PSIs, potential risk-adjustment covariates indicate whether the discharge record meets the technical specification for sex, age, MDRGs, MDCs, and a list of 25 comorbidity variables from the HCUP project.

For the IQIs, potential risk-adjustment covariates indicate whether the discharge record meets the technical specification for sex, age, APR-DRGs, and risk-of-mortality subclass (minor, moderate, major, extreme) that are used as covariates in the risk-adjustment model.

For the PDIs, potential risk-adjustment covariates indicate whether the discharge record meets the technical specification for birth weight, sex, age in days, age in years, MDRG, at least 1 of 46 CCS comorbidities, and some indicator-specific risk categories that are used as covariates in the risk-adjustment model.

For the PQIs, potential risk-adjustment covariates indicate whether the discharge record meets the technical specification for sex, age in 5-year groups, and poverty category that are used as covariates in the risk-adjustment model.

Covariates are coded for each discharge record on the basis of the data elements, data values, and logic described in the technical specifications and the appendices of the risk-adjustment coefficient tables. For a given covariate, if the discharge meets the technical specification for that covariate, a value of 1 is assigned to the discharge-level covariate data element. Otherwise, a value of 0 is assigned to the discharge-level covariate data element.

Select Model Covariates

On the basis of cross tabulations between each covariate and the outcome of interest, only those covariates with at least 30 cases with the outcome of interest are retained. The omitted covariate within mutually exclusive categories is the reference group for those categories. Reference categories are usually (1) the most common and/or (2) the least risk, or (3) the median category. The choice of omitted reference category does affect how one might describe the parameter coefficients in words, but it does not affect predicted probabilities or model performance.

Variables for inclusion in the final risk adjustment models are selected by bootstrapped regression. Up to 1,000 bootstrap samples with replacement were selected from the reference population. The number of bootstrap replicates was limited for some indicators because of longer run times for each replicate. For each replicate sample, a logistic model was fit with the QI T-flag as the dependent variable. Backwards stepwise selection was used with a threshold of $p < 0.1$ to keep a variable in the model from the candidate pool of covariates. Variables kept in more than 60 percent of the replicates were retained in the final model. Variables for an age-sex interaction are defined as compulsory for the model.

For the area-level indicators, the models use the complete set of covariates for sex, age in 5-year age groups, an interaction with sex * age. There is also an optional set of covariates for poverty category based on the county of patient residence.

The final multivariable model parameters are published on the AHRQ Web site in Risk Adjustment Coefficient Tables. (See [links](#) in the Overview chapter.)

Estimate the Models

When possible, provider-level models are estimated using GEEs to account for within-hospital correlation. These models are run with PROC GENMOD and use a logit link with an exchangeable covariance matrix. If the GEE model does not converge, then a more logistic regression model is fit (i.e., PROC LOGISTIC) that ignores that extra correlation. Whether the model is a GEE may be inferred by the final column in the .CSV file for the QI. Area-level indicators use logistic models.

Calculate Rates

After the new risk-adjustment models are fit, PROC SCORE is run on the data to calculate expected values so that observed rates may be calculated for the reference population. Reference population rates and signal variances are calculated both ignoring POA altogether and with POA as recorded. These rates are stored in .TXT files that are part of the SAS AHRQ QI software package. The rates and variances are entered directly into WinQI program code and do not appear as separate files in the WinQI package. Updating the risk-adjustment .CSV files and the population rate and signal variance .TXT files are a substantial milestone in the annual update process.

Evaluate Models

Two desirable qualities of risk-adjustment models are that they discriminate well between discharge records that experience the outcome of interest and those that do not and that they are well calibrated, predicting that the outcome will occur in approximately the right proportions, over a wide range of predicted probability.

Discrimination

One common scalar measure of logistic regression discrimination is the c-statistic. This may be calculated by computing the area under the Receiver Operating Characteristic (ROC) curve. Alternatively, it may be calculated by forming every possible pair in a dataset in which one member of the pair is a discharge with the outcome of interest and the other member is a discharge

without the outcome of interest. The c-statistic is the proportion of such pairs in which the predicted probability for the member with the outcome of interest is higher than the predicted probability for the other record. Pairs with tied probabilities each contribute one-half to the numerator and denominator of the proportion. A c-statistic of 0.5 is the same discrimination performance as flipping a coin. A c-statistic of 1.0 indicates perfect discrimination. Hosmer and Lemeshow (2000, p. 162) have coined three widely adopted labels for discrimination performance based on the c-statistic:

- $0.70 \leq \text{c-statistic} < 0.80$ indicates acceptable discrimination
- $0.80 \leq \text{c-statistic} < 0.90$ indicates excellent discrimination
- $0.90 \leq \text{c-statistic}$ indicates outstanding discrimination

The c-statistics for the AHRQ QI risk-adjustment models are published in on the AHRQ QI Web site in the Risk Adjustment Coefficient Tables. (See [links](#) in the Overview chapter.)

Calibration

Calibration often is described by sorting the dataset on the basis of predicted probability and dividing it into deciles of risk. It is meaningful to compare the proportion of records in each decile that were observed to have the outcome of interest with the proportion of records that are expected to have that outcome. Hosmer and Lemeshow's (1980) logistic regression goodness-of-fit statistic is based on a chi-square test statistic calculated using the observed and expected counts across the 10 deciles. Unfortunately, that statistic always rejects the null hypothesis good calibration when the number of observations is large, as is the case with the AHRQ QI reference population. Although the test statistic and its p-value are not informative for these models, the models are sometimes characterized by publishing or plotting the observed and expected counts in the 10 deciles of risk.

Chapter 12. Other Files Referenced in AHRQ QI Software

The AHRQ QI software uses several other files or datasets that are updated periodically. This chapter lists those and either describes the methods used to generate them or references other stand-alone documents that do so.

Population Reference File

The file that contains stratified population counts by county and MSA is crucial for calculating the denominators of the area-level measures. That file and the method to construct it are described in a file titled “AHRQ QI Population File Documentation” on the AHRQ Web site: (<http://www.qualityindicators.ahrq.gov/software/SAS.aspx>)

Condition-Specific Population File

The AHRQ QI program includes ongoing research into options for estimating condition-specific denominators. At this time, the only condition-specific denominators are related to diabetes. There is a file name QICTYC14.TXT that is included with AHRQ PQI module. That file was calculated using the following steps:

1. Use the census population denominator reference file to estimate the 2014 population for each combination of State and age category. In the QI software, age categories are coded as follows:

VALUE AGECCAT
0 = '00 to 17'
1 = '18 to 44'
2 = '45 to 64'
3 = '65 to 74'
4 = '75+'
2. Obtain the latest diabetes prevalence figures broken out by State and age category from the Centers for Disease Control at <http://gis.cdc.gov/grasp/diabetes/DiabetesAtlas.html>
3. Apply the diabetes proportions to the populations to estimate the number of adults in each State in each of the four age categories who would have diabetes in 2014 (population data from 2014 and proportion data from 2012).

References

Hosmer DW, Lemeshow S. Confidence interval estimates of an index of quality performance mcelbased on logistic regression. *Statistics in Med.* 1995;14(19):2161-72.

Iezzoni, Lisa, Ed. *Risk Adjustment for Measuring Health Care Outcomes*, 4th ed. Chicago: Health Administration Press; 2013.

Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika.* 1986 April;73(1):13-22.

Luft HS, Brown BW Jr. Calculating the probability of rare events: why settle for an approximation? *Health Serv Res.* 1993;28(4):419-39.

McClellan M, Staiger D. The quality of health care providers. NBER Working Paper #7327. Cambridge, MA: National Bureau of Economic Research; 1999.
<http://www.nber.org/papers/w7327>. Accessed November 10, 2015.

Morris, CN. Parametric empirical Bayes inference: theory and applications. *J Am Statistical Assoc.* 1983 Mar;78(381):47-55.

Phibbs CS, Baker LC, Caughey AB, et al. Level and volume of neonatal intensive care and mortality in very-low-birth-weight infants. *N Engl J Med.* 2007;356(21):2165-75 & Supplement.

Appendix A. Table of AHRQ QI Risk-Adjustment / POA

Appendix Table A.1 denotes which Agency for Healthcare Research and Quality (AHRQ) Quality Indicators (QIs) are risk adjusted and which use Present on Admission (POA) data and for what purpose (i.e., for technical specifications or risk adjustment).

An entry of *X* in the column titled Calculate Risk Adjusted Rate means that the indicator is risk adjusted using PROC SCORE in SAS with coefficients from the risk-adjustment models estimated using generalized estimating equation (GEE) or LOGISTIC models (Liang and Zieger, 1986).

An *X* in the column marked Technical Specifications means that the indicator has an exclusion that explicitly references the POA data element. When a discharge record is missing the DX_POA data element, the Q flag will be set to missing (.) and the software will ignore it.

An *X* in the column marked Risk Adjustment means that the risk adjustment logistic regression model includes covariates for conditions that are comorbidities if they are POA and are complications if they are not POA. When the discharge record is missing the DX_POA data element, the risk adjustment model will treat the condition as if it were a complication that was not POA.

See [Chapter 9](#) for additional details on risk adjustment.

Appendix Table A.1. AHRQ QI Risk-Adjustment and Uses of POA

Indicator	Calculate Risk-Adjusted Rate	Use POA? Technical Specifications	Use POA? Risk Adjustment
Inpatient Quality Indicators			
IQI 01 - Esophageal Resection Volume			
IQI 02 - Pancreatic Resection Volume			
IQI 04 - Abdominal Aortic Aneurysm (AAA) Repair Volume			
IQI 05 - Coronary Artery Bypass Graft (CABG) Volume			
IQI 06 - Percutaneous Coronary Intervention (PCI) Volume			
IQI 07 - Carotid Endarterectomy Volume			
IQI 08 - Esophageal Resection Mortality Rate			•
IQI 09 - Pancreatic Resection Mortality Rate	•		•
IQI 11 - Abdominal Aortic Aneurysm (AAA) Repair Mortality Rate	•		•
IQI 12 - Coronary Artery Bypass Graft (CABG) Mortality Rate	•		•
IQI 13 - Craniotomy Mortality Rate	•		•
IQI 14 - Hip Replacement Mortality Rate	•		•

Indicator	Calculate Risk-Adjusted Rate	Use POA? Technical Specifications	Use POA? Risk Adjustment
IQI 15 - Acute Myocardial Infarction (AMI) Mortality Rate	●		●
IQI 16 - Heart Failure Mortality Rate	●		●
IQI 17 - Acute Stroke Mortality Rate	●		●
IQI 18 - Gastrointestinal Hemorrhage Mortality Rate	●		●
IQI 19 - Hip Fracture Mortality Rate	●		●
IQI 20 - Pneumonia Mortality Rate	●		●
IQI 21 - Cesarean Delivery Rate, Uncomplicated			
IQI 22 - Vaginal Birth After Cesarean (VBAC) Delivery Rate, Uncomplicated			
IQI 23 - Laparoscopic Cholecystectomy Rate			
IQI 24 - Incidental Appendectomy in the Elderly Rate			
IQI 25 - Bi-lateral Cardiac Catheterization Rate			
IQI 26 - Coronary Artery Bypass Graft (CABG) Rate	●		
IQI 27 - Percutaneous Coronary Intervention (PCI) Rate	●		
IQI 28 - Hysterectomy Rate	●		
IQI 29 - Laminectomy or Spinal Fusion Rate	●		
IQI 30 - Percutaneous Coronary Intervention (PCI) Mortality Rate	●		●
IQI 31 - Carotid Endarterectomy Mortality Rate	●		●
IQI 32 - Acute Myocardial Infarction (AMI) Mortality Rate, Without Transfer Cases	●		●
IQI 33 - Primary Cesarean Delivery Rate, Uncomplicated			
Patient Safety Indicators			
PSI 02 - Death Rate in Low-Mortality Diagnosis Related Groups (DRGs)	●		●
PSI 03 - Pressure Ulcer Rate	●	●	●
PSI 04 - Death Rate among Surgical Inpatients with Serious Treatable Complications	●		●
PSI 05 - Retained Surgical Item or Unretrieved Device Fragment Count		●	
PSI 06 - Iatrogenic Pneumothorax Rate	●	●	●
PSI 07 - Central Venous Catheter-Related Blood Stream Infection Rate	●	●	●
PSI 08 - Postoperative Hip Fracture Rate	●	●	●
PSI 09 - Perioperative Hemorrhage or Hematoma Rate	●	●	●
PSI 10 - Postoperative Physiologic and Metabolic Derangement Rate	●	●	●
PSI 11 - Postoperative Respiratory Failure Rate	●	●	●
PSI 12 - Perioperative Pulmonary Embolism or Deep Vein Thrombosis Rate	●	●	●
PSI 13 - Postoperative Sepsis Rate	●	●	●

Indicator	Calculate Risk-Adjusted Rate	Use POA? Technical Specifications	Use POA? Risk Adjustment
PSI 14 - Postoperative Wound Dehiscence Rate	●		●
PSI 15 - Accidental Puncture or Laceration Rate	●	●	●
PSI 16 - Transfusion Reaction Count		●	
PSI 17 - Birth Trauma Rate – Injury to Neonate			
PSI 18 - Obstetric Trauma Rate – Vaginal Delivery With Instrument			
PSI 19 - Obstetric Trauma Rate – Vaginal Delivery Without Instrument			
Pediatric Quality Indicators			
PDI 01 - Accidental Puncture or Laceration Rate	●	●	●
PDI 02 - Pressure Ulcer Rate	●	●	●
PDI 03 - Retained Surgical Item or Unretrieved Device Fragment Count		●	
PDI 05 - Iatrogenic Pneumothorax Rate	●	●	●
PDI 06 - RACHS-1 Pediatric Heart Surgery Mortality Rate	●		●
PDI 07 - RACHS-1 Pediatric Heart Surgery Volume			
PDI 08 - Perioperative Hemorrhage or Hematoma Rate	●	●	●
PDI 09 - Postoperative Respiratory Failure Rate	●	●	●
PDI 10 - Postoperative Sepsis Rate	●	●	●
PDI 11 - Postoperative Wound Dehiscence Rate			●
PDI 12 - Central Venous Catheter-Related Blood Stream Infection Rate	●	●	●
PDI 13 - Transfusion Reaction Count		●	
PDI 14 – Asthma Admission Rate	●		
PDI 15 – Diabetes Short-Term Complications Admission Rate	●		
PDI 16 – Gastroenteritis Admission Rate	●		
PDI 17 – Perforated Appendix Admission Rate	●		
PDI 18 – Urinary Tract Infection Admission Rate	●		
NQI 01 - Neonatal Iatrogenic Pneumothorax Rate		●	●
NQI 02 - Neonatal Mortality Rate	●		●
NQI 03 - Neonatal Blood Stream Infection Rate	●	●	●
Prevention Quality Indicators			
PQI 01 - Diabetes Short-Term Complications Admission Rate	●		
PQI 02 - Perforated Appendix Admission Rate	●		
PQI 03 - Diabetes Long-Term Complications Admission Rate	●		
PQI 05 - Chronic Obstructive Pulmonary Disease (COPD) or Asthma in Older Adults Admission Rate	●		
PQI 07 - Hypertension Admission Rate	●		
PQI 08 - Heart Failure Admission Rate	●		
PQI 09 - Low Birth Weight Rate	●		
PQI 10 - Dehydration Admission Rate	●		

Indicator	Calculate Risk-Adjusted Rate	Use POA? Technical Specifications	Use POA? Risk Adjustment
PQI 11 - Bacterial Pneumonia Admission Rate	●		
PQI 12 - Urinary Tract Infection Admission Rate	●		
PQI 13 - Angina Without Procedure Admission Rate	●		
PQI 14 - Uncontrolled Diabetes Admission Rate	●		
PQI 15 - Asthma in Younger Adults Admission Rate	●		
PQI 16 - Lower-Extremity Amputation among Patients with Diabetes Rate	●		

Appendix B. Table of AHRQ QI Provider-Level Risk-Adjustment Covariates

The categories highlighted in blue are mutually exclusive and exhaustive, meaning that every discharge is assigned a value of 1 for one and only one covariate and there must be an omitted covariate (usually the most common or the least risk). If covariates within a highlighted category are excluded because $N < 30$ or $p < 0.05$, then the covariate is combined with another along the risk gradient. For example, combine birth weight 500–999g with 1000–1499g, ages 18–24 years with ages 25–29 years, or ROM subclass 4 with ROM subclass 3.

Appendix Table B.1. AHRQ QI Risk-Adjustment Covariates for Provider-Level Indicators

Category	Mutually Exclusive	IQI	PSI	PDI	NQI
Demographics		Sex	Sex	Sex	Sex
		Age (5-year age groups)	Age (5-year age groups)	Birth weight (500g groups) Age in days (90 days–1 year) Age in years (1 year+)	Birth weight (500g groups)
Severity of Illness	DRGs pool into MDCs	APR-DRG MDCs	Modified MS-DRG* MDCs	Modified MS-DRG ^a MDCs	Modified MS-DRG ^a MDCs
Comorbidities		APR-DRG Risk of mortality subclass (1 – minor; 2 – moderate; 3 – major; 4 – extreme)	AHRQ Comorbidities	AHRQ Clinical Classification Software	Congenital anomalies

Category	Mutually Exclusive	IQI	PSI	PDI	NQI
Other		Transfer-in status Point of Origin status	Transfer-in status Point of Origin status Days to Procedure status	Transfer-in status Point of Origin status Days to Procedure status Indicator-specific risk stratifiers	Transfer-in status Point of Origin status Days to Procedure status

Abbreviations: AHRQ, Agency for Healthcare Research and Quality; APR-DRG, all patient refined diagnostic related group; IQI, Inpatient Quality Indicator; MDC, major diagnostic category; MS-DRG, Medicare severity diagnostic related group; NQI, Neonatal Quality Indicator; PDI, Pediatric Quality Indicator; PSI, Patient Safety Indicator; QI, Quality Indicator;

^a Prior to October 1, 2007, use CMS-DRGs; highlighted categories are mutually exclusive with an omitted covariate.

Appendix C. External Resources Referenced by the AHRQ QI

This appendix includes references to resource used in the, the Agency for Healthcare Research and Quality (AHRQ) Quality Indicator (QI) program. This information is not specifically statistical in nature but does inform and affect the methods described in the main body of the document.

A. Fiscal Year Coding Updates

Each fiscal year, there are new ICD-10-CM and MS-DRG codes and revisions to existing codes. These changes are effective on October 1. For example, version 33 (fiscal year 2016) codes were effective October 1, 2015, and were incorporated in the version 6.0 release of the QI software. Diagnosis and procedure codes are used in the numerator and denominator specifications for the Patient Safety Indicators (PSIs), Prevention Quality Indicators (PQIs), Pediatric Quality Indicators (PDIs), and Inpatient Quality Indicators (IQIs). ICD-10-CM procedure codes affect the Centers for Medicare & Medicaid Services (CMS) classification of “major operating room procedure” for postoperative PSIs and PDIs. Another use of ICD-10-CM is in risk stratification used in the AHRQ Comorbidity Software, AHRQ’s Clinical Classification System, and 3M’s all patient refined diagnosis related groups (APR-DRGs). Diagnosis codes are maintained by the Centers for Disease Control and Prevention (CDC) National Center for Health Statistics (NCHS). CMS maintains procedure and Medicare severity diagnosis related codes (MS-DRG) codes. The activities of both agencies are conducted jointly through the ICD-10-CM Coordination and Maintenance Committee (the Committee). The Committee meets in September and March to consider proposals for new codes and revisions to existing codes.

The Committee has implemented a partial freeze of the ICD-9-CM and ICD-10-CM/PCS codes in preparation for the implementation of ICD-10 codes on October 1, 2013. As a result, the last regular, annual updates to both ICD-9-CM and ICD-10-CM/PCS codes were made on October 1, 2011 (fiscal year 2012). Following October 1, 2012, only limited coding updates were made to both the ICD-9-CM and ICD-10-CM/PCS codes to capture new technologies and diseases.

Information on ICD-10-CM coding updates is located on both the NCHS (<http://www.cdc.gov/nchs/icd/icd10cm.htm>) and CMS (<http://www.cms.gov/ICD10>) Web sites.

3M maintains APR-DRG codes.

A.1 ICD-10-CM Coding Updates and Coding Guidelines

Information on ICD-10-CM coding updates is located on the NCHS and CMS Web sites:

- <http://www.cdc.gov/nchs/icd/icd10cm.htm>
- http://www.cdc.gov/nchs/data/icd/10cmguidelines_2016_Final.pdf

3M APR-DRG

The AHRQ QI software includes a limited license group that facilitates the use of the QIs and APR-DRGs together. The grouper is updated as new versions are available. 3M currently releases a new version each fiscal year.

B. Related Software Maintained by HCUP at AHRQ

The AHRQ QI software uses other AHRQ software as components of the indicator specifications or risk-adjustment covariate specifications. These software components also are updated annually to reflect coding changes. The AHRQ QI support team does not review these changes independently; rather the coding changes are implemented without further review.

B.1 Comorbidity Software

Variables created by the AHRQ Comorbidity Software are used in the AHRQ PSI risk adjustment models. There are two editions of the comorbidity software. The first edition uses CMS-DRGs, and the second edition uses MS-DRGs. The comorbidity software has its own version numbering system. The first edition is version 3.4 and earlier; the second edition is version 3.5 and later (see <http://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp>).

B.2 Clinical Classification Software (CCS)

The CCS for ICD-10-CM is a diagnosis and procedure categorization scheme that collapses individual codes into a smaller number of clinically meaningful categories. The AHRQ QI uses the single-level edition of the CCS for diagnoses and procedures. The software consists of a SAS formats program (see <http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp>).

B.3 Procedure Classes

The procedure classes assign ICD-10-CM procedure codes to one of four categories:

- Minor Diagnostic – Nonoperating room procedures that are diagnostic
- Minor Therapeutic – Nonoperating room procedures that are therapeutic
- Major Diagnostic – All procedures considered valid operating room procedures by the DRG grouper and that are performed for diagnostic reasons
- Major Therapeutic – All procedures considered valid operating room procedures by the DRG grouper and that are performed for therapeutic reasons

(See <http://www.hcup-us.ahrq.gov/toolssoftware/procedure/procedure.jsp>.)

C. Related Classifications Maintained by the AHRQ QI Support Team

The AHRQ QI software also uses other classifications as a component of the indicator specification or risk-adjustment covariate specification. These classification components are updated annually to reflect coding changes. The classifications include the modified DRGs (MDRGs), birth weight (BWHTCAT), Congenital Anomalies (CONGCAT), and indicator-specification stratifications for the PDIs (HPPD01, GPPD02, GPPD10, HPPD10 and GPPD12).

C.1 Modified DRGs (MDRGs)

The purpose of the MDRG is to pool MS-DRGs with and without CCs and MCCs. A new MS-DRG code either divides an existing MS-DRG into sub-MS-DRGs or re-assigns cases from multiple existing MS-DRGs. The MDRG is a 4-digit code. The first 2 digits are the major diagnosis category (MDC), and the second 2 digits are a sequence number (e.g., 01-04) within the MDC.

C.2 Birth Weight (BWHTCAT)

BWHTCAT in 250g increments are defined by ICD-9-CM codes. Occasionally, new codes are derived from existing codes.

C.3 Congenital Anomalies (CONGCAT)

CONGCAT for gastrointestinal, genitourinary, central nervous system, pulmonary, cardiovascular, skeletal, chromosomal syndromes, and selected other congenital anomalies are defined by ICD-10-CM codes (original source Phibbs et al.). Occasionally, new codes are derived from existing codes.

C.4 Indicator-Specific

Some PDIs have classifications used in stratification and as covariates in risk adjustment. These classifications are procedure type risk category (HPPD01), pressure ulcer risk category (GPPD02), wound class procedure type (GPPD10), immune-compromised risk category (HPPD10), and bloodstream infection risk category (GPPD12). Occasionally new codes are derived from existing codes.

Updating the indicator-specific classifications consists of the following steps:

1. Identify the relevant ICD-10-CM coding updates that pertain to the definition of the classifications.
2. Update the specifications, appendix, and change log for the relevant AHRQ QIs.
3. Implement any changes in the AHRQ QI software.

D. Risk Adjustment for Congenital Heart Surgery (RACHS-1) Software

RACHS-1 is a type of specification (the numerator and denominator inclusion and exclusion rules). The Pediatric Heart Surgery Mortality (PDI 06) measure uses the RACHS-1 software to assign pediatric heart surgery cases to risk strata depending on the type of surgery (HPPD06). The stratification occurs upon running the RACHS-1 syntax, which is embedded in the software. Children's Hospital in Boston maintains the RACHS-1 software on an ad hoc basis (see <http://www.ncbi.nlm.nih.gov/pubmed/15283367>).