

## **Uncontrolled Diabetes Admission Rate (PQI 14)**

### **Table of Contents for Supporting Materials**

(NOTE: majority of materials found at: [www.qualityindicators.ahrq.gov](http://www.qualityindicators.ahrq.gov))

#### Technical specification of the indicator

1. Technical specifications
2. PQI appendices
3. PQI Parameter Estimates
4. PQI coding updates

#### Methodology

5. AHRQ QI Empirical methods
6. Population files

#### Scientific Acceptability

7. Analytic Template
8. HCUPnet Results, 2011

# Uncontrolled Diabetes Admission Rate

## Technical Specifications

### Prevention Quality Indicators #14 (PQI #14)

AHRQ Quality Indicators™, Version 4.5, May 2013

Area-Level Indicator

Type of Score: Rate

### Description

Admissions for a principal diagnosis of diabetes without mention of short-term (ketoacidosis, hyperosmolarity, or coma) or long-term (renal, eye, neurological, circulatory, or other unspecified) complications per 100,000 population, ages 18 years and older. Excludes obstetric admissions and transfers from other institutions.

*[NOTE: The software provides the rate per population. However, common practice reports the measure as per 100,000 population. The user must multiply the rate obtained from the software by 100,000 to report admissions per 100,000 population.]*

### Numerator

Discharges, for patients ages 18 years and older, with a principal ICD-9-CM diagnosis code for uncontrolled diabetes without mention of a short-term or long-term complication.

*[NOTE: By definition, discharges with a principal diagnosis of uncontrolled diabetes without mention of short-term or long-term complications are precluded from an assignment of MDC 14 by grouper software. Thus, obstetric discharges should not be considered in the PQI rate, though the AHRQ QI™ software does not explicitly exclude obstetric cases.]*

#### ICD-9-CM Uncontrolled diabetes without mention of a short-term or long-term complication diagnosis codes:

25002    DMII WO CMP UNCNTRLD

25003    DMI WO CMP UNCNTRLD

Exclude cases:

- transfer from a hospital (different facility)
- transfer from a Skilled Nursing Facility (SNF) or Intermediate Care Facility (ICF)
- transfer from another health care facility
- with missing gender (SEX=missing), age (AGE=missing), quarter (DQTR=missing), year (YEAR=missing), principal diagnosis (DX1=missing), or county (PSTCO=missing)

See *Prevention Quality Indicators Appendices*:

- Appendix A – Admission Codes for Transfers

## **Denominator**

Population ages 18 years and older in metropolitan area<sup>†</sup> or county. Discharges in the numerator are assigned to the denominator based on the metropolitan area or county of the patient residence, not the metropolitan area or county of the hospital where the discharge occurred.<sup>‡</sup> May be combined with diabetes short-term complications as a single indicator as a simple sum of the rates to form the Health People 2010 indicator (note that the AHRQ QI™ excludes transfers to avoid double counting cases).

---

<sup>†</sup> 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 (CMSAs), and primary metropolitan statistical areas (PMSAs). In addition, “area” could refer to either 1) FIPS county, 2) modified FIPS county, 3) 1999 OMB Metropolitan Statistical Area, or 4) 2003 OMB Metropolitan Statistical Area. Micropolitan Statistical Areas are not used in the QI software.

<sup>‡</sup> The denominator can be specified with the diabetic population only and calculated with the SAS QI software through the condition-specific denominator at the state-level feature.



# **TECHNICAL SPECIFICATIONS: PREVENTION QUALITY INDICATORS APPENDICES**

## **Version 4.5**

**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. HHSA290201200001C**

**Prepared by:**

**Battelle**  
505 King Avenue  
Columbus, OH 43201

**May 2013**

## Table of Contents

Appendix A – Admission Codes for Transfers .....	1
Appendix B – Cardiac Procedure Codes .....	2
Appendix C – Immunocompromised State Diagnosis and Procedure Codes .....	4
Appendix D – Definitions of Neonate, Newborn, Normal Newborn, and Outborn .....	6

## **Appendix A – Admission Codes for Transfers**

### **SID ASOURCE Codes**

- 2 - Another hospital
- 3 - Another facility, including long-term care

### **POINTOFORIGINUB04 Codes**

- 4 - Transfer from a hospital
- 5 - Transfer from a Skilled Nursing Facility (SNF) or Intermediate Care Facility (ICF)
- 6 - Transfer from another health care facility

If Admission Type is newborn (ATYPE=4), POINTOFORIGINUB04 codes are as follows:

- 5 - Born inside this hospital
- 6 - Born outside of this hospital

## Appendix B – Cardiac Procedure Codes

### ICD-9-CM Cardiac procedure codes<sup>1</sup>:

0050	IMPL CRT PACEMAKER SYS	3582	TOTAL REPAIR OF TAPVC
0051	IMPL CRT DEFIBRILLAT	3583	TOT REP TRUNCUS ARTERIOS
0052	IMP/REP LEAD LF VEN SYS	3584	TOT COR TRANSPOS GRT VES
0053	IMP/REP CRT PACEMKR GEN	3591	INTERAT VEN RETRN TRANSP
0054	IMP/REP CRT DEFIB GENAT	3592	CONDUIT RT VENT-PUL ART
0056	INS/REP SENS-CRD/VSL MTR	3593	CONDUIT LEFT VENTR-AORTA
0057	IMP/REP SUBCUE CARD DEV	3594	CONDUIT ARTIUM-PULM ART
0066	PTCA	3595	HEART REPAIR REVISION
1751	IMPLANT CCM,TOTAL SYSTEM	3596	PERC BALLOON VALVULOPLASTY
1752	IMPLANT CCM PULSE GENRTR	3597	PERC MTRL VLV REPR W IMP
1755	TRANSLUM COR ATHERECTOMY	3598	OTHER HEART SEPTA OPS
3500	CLOSED VALVOTOMY NOS	3599	OTHER HEART VALVE OPS
3501	CLOSED AORTIC VALVOTOMY	3601	<i>PTCA-1 VESSEL W/O AGENT</i>
3502	CLOSED MITRAL VALVOTOMY	3602	<i>PTCA-1 VESSEL WITH AGNT</i>
3503	CLOSED PULMON VALVOTOMY	3603	OPEN CORONRY ANGIOPLASTY
3504	CLOSED TRICUSP VALVOTOMY	3604	INTRACORONRY THROMB INFUS
3505	ENDOVAS REPL AORTC VALVE	3605	<i>PTCA-MULTIPLE VESSEL</i>
3506	TRNSAPCL REP AORTC VALVE	3606	INS NONDRUG ELUT COR ST
3507	ENDOVAS REPL PULM VALVE	3607	INS DRUG-ELUT CORONRY ST
3508	TRNSAPCL REPL PULM VALVE	3609	REM OF COR ART OBSTR NEC
3509	ENDOVAS REPL UNS HRT VLV	3610	AORTOCORONARY BYPASS NOS
3510	OPEN VALVULOPLASTY NOS	3611	AORTOCOR BYPAS-1 COR ART
3511	OPN AORTIC VALVULOPLASTY	3612	AORTOCOR BYPAS-2 COR ART
3512	OPN MITRAL VALVULOPLASTY	3613	AORTOCOR BYPAS-3 COR ART
3513	OPN PULMON VALVULOPLASTY	3614	AORTCOR BYPAS-4+ COR ART
3514	OPN TRICUS VALVULOPLASTY	3615	1 INT MAM-COR ART BYPASS
3520	OPN/OTH REP HRT VLV NOS	3616	2 INT MAM-COR ART BYPASS
3521	OPN/OTH REP AORT VLV-TIS	3617	ABD-CORON ARTERY BYPASS
3522	OPN/OTH REP AORTIC VALVE	3619	HRT REVAS BYPS ANAS NEC
3523	OPN/OTH REP MTRL VLV-TIS	362	ARTERIAL IMPLANT REVASC
3524	OPN/OTH REP MITRAL VALVE	363	<i>OTH HEART REVASCULAR</i>
3525	OPN/OTH REP PULM VLV-TIS	3631	OPEN CHEST TRANS REVASC
3526	OPN/OTH REPL PUL VALVE	3632	OTH TRANSMYO REVASCULAR
3527	OPN/OTH REP TCSPD VLV-TS	3633	ENDO TRANSMYO REVASCULAR
3528	OPN/OTH REPL TCSPD VALVE	3634	PERC TRANSMYO REVASCULAR
3531	PAPILLARY MUSCLE OPS	3639	OTH HEART REVASULAR
3532	CHORDAE TENDINEAE OPS	3691	CORON VESS ANEURYSM REP
3533	ANNULOPLASTY	3699	HEART VESSEL OP NEC
3534	INFUNDIBULECTOMY	3731	PERICARDIECTOMY
3535	TRABECUL CARNEAE CORD OP	3732	HEART ANEURYSM EXCISION
3539	TISS ADJ TO VALV OPS NEC	3733	EXC/DEST HRT LESION OPEN
3541	ENLARGE EXISTING SEP DEF	3734	EXC/DEST HRT LES OTHER
3542	CREATE SEPTAL DEFECT	3735	PARTIAL VENTRICULECTOMY
3550	PROSTH REP HRT SEPTA NOS	3736	EXC,DESTRCT,EXCLUS LAA
3551	PROS REP ATRIAL DEF-OPN	3737	EXC/DEST HRT LES, THRSPC
3552	PROS REPAIR ATRIA DEF-CL	3741	IMPL CARDIAC SUPPORT DEV
3553	PROS REP VENTRIC DEF-OPN	375	<i>HEART TRANSPLANTATION</i>
3554	PROS REP ENDOCAR CUSHION	3751	HEART TRANPLANTATION
3555	PROS REP VENTRC DEF-CLOS	3752	IMP TOT INT BI HT RP SYS
3560	GRFT REPAIR HRT SEPT NOS	3753	REPL/REP THR UNT TOT HRT
3561	GRAFT REPAIR ATRIAL DEF	3754	REPL/REP OTH TOT HRT SYS
3562	GRAFT REPAIR VENTRIC DEF	3755	REM INT BIVENT HRT SYS
3563	GRFT REP ENDOCAR CUSHION	3760	IMP BIVN EXT HRT AST SYS
3570	HEART SEPTA REPAIR NOS	3761	PULSATION BALLOON IMPLAN
3571	ATRIA SEPTA DEF REP NEC	3762	INSRT NON-IMPL CIRC DEV
3572	VENTR SEPTA DEF REP NEC	3763	REPAIR HEART ASSIST SYS
3573	ENDOCAR CUSHION REP NEC	3764	REMVE EXT HRT ASSIST SYS
3581	TOT REPAIR TETRAL FALLOT	3765	IMP VENT EXT HRT AST SYS

3766	IMPLANTABLE HRT ASSIST	3782	INT INSERT 1-CHAM, RATE
3770	INT INSERT PACEMAK LEAD	3783	INT INSERT DUAL-CHAM DEV
3771	INT INSERT LEAD IN VENT	3785	REPL PACEM W 1-CHAM, NON
3772	INT INSER LEAD ATRI-VENT	3786	REPL PACEM 1-CHAM, RATE
3773	INT INSER LEAD IN ATRIUM	3787	REPL PACEM W DUAL-CHAM
3774	INT OR REPL LEAD EPICAR	3789	REVISE OR REMOVE PACEMAK
3775	REVISION OF LEAD	3794	IMPLT/REPL CARDDEFIB TOT
3776	REPL TV ATRI-VENT LEAD	3795	IMPLT CARDIODEFIB LEADS
3777	REMOVAL OF LEAD W/O REPL	3796	IMPLT CARDIODEFIB GENRATR
3778	INSER TEAM PACEMAKER SYS	3797	REPL CARDIODEFIB LEADS
3779	REV/RELOC CARD DEV POCKT	3798	REPL CARDIODEFIB GENRATR
3780	INT OR REPL PERM PACEMKR	3826	INSRT PRSR SNSR W/O LEAD
3781	INT INSERT 1-CHAM, NON		

<sup>1</sup>The procedure or diagnosis codes are continuously updated. The current list of ICD-9-CM codes is valid for October 2012 through September 2013. Italicized codes are not active in Fiscal Year 2013.



## Appendix C – Immunocompromised State Diagnosis and Procedure Codes

### ICD-9-CM Immunocompromised state diagnosis codes<sup>1</sup>:

042	HUMAN IMMUNO VIRUS DIS	2882	GENETIC ANOMALY LEUKOCYT
1363	PNEUMOCYSTOSIS	2884	HEMOPHAGOCYTIC SYNDROMES
1992	MALIG NEOPL-TRANSP ORGAN	28850	LEUKOCYTOPENIA NOS
23873	HI GRDE MYELOYDYS SYN LES	28851	LYMPHOCYTOPENIA
23876	MYELOFI W MYELO METAPLAS	28859	DECREASED WBC COUNT NEC
23877	POST TP LYMPHPROLIF DIS	28953	NEUTROPENIC SPLENOMEGALY
23879	LYMPH/HEMATPOITC TIS NEC	28983	MYELOFIBROSIS
260	KWASHIORKOR	40301	MAL HYP KID W CR KID V
261	NUTRITIONAL MARASMUS	40311	BEN HYP KID W CR KID V
262	OTH SEVERE MALNUTRITION	40391	HYP KID NOS W CR KID V
27900	HYPOGAMMAGLOBULINEM NOS	40402	MAL HY HT/KD ST V W/O HF
27901	SELECTIVE IGA IMMUNODEF	40403	MAL HYP HT/KD STG V W HF
27902	SELECTIVE IGM IMMUNODEF	40412	BEN HY HT/KD ST V W/O HF
27903	SELECTIVE IG DEFIC NEC	40413	BEN HYP HT/KD STG V W HF
27904	CONG HYPOGAMMAGLOBULINEM	40492	HY HT/KD NOS ST V W/O HF
27905	IMMUNODEFIC W HYPER-IGM	40493	HYP HT/KD NOS ST V W HF
27906	COMMON VARIABL IMMUNODEF	5793	INTEST POSTOP NONABSORB
27909	HUMORAL IMMUNITY DEF NEC	585	<i>CHRONIC KIDNEY DISEASE</i>
27910	IMMUNDEF T-CELL DEF NOS	5855	CHRON KIDNEY DIS STAGE V
27911	DIGEORGE'S SYNDROME	5856	END STAGE RENAL DISEASE
27912	WISKOTT-ALDRICH SYNDROME	9968	<i>COMPLICATIONS OF TRANSPLANTED ORGAN</i>
27913	NEZELOF'S SYNDROME	99680	COMP ORGAN TRANSPLNT NOS
27919	DEFIC CELL IMMUNITY NOS	99681	COMPL KIDNEY TRANSPLANT
2792	COMBINED IMMUNITY DEFIC	99682	COMPL LIVER TRANSPLANT
2793	IMMUNITY DEFICIENCY NOS	99683	COMPL HEART TRANSPLANT
2794	<i>AUTOIMMUNE DISEASE, NOT ELSEWHERE CLASSIFIED</i>	99684	COMPL LUNG TRANSPLANT
27941	AUTOIMMUN LYMPHPROF SYND	99685	COMPL MARROW TRANSPLANT
27949	AUTOIMMUNE DISEASE NEC	99686	COMPL PANCREAS TRANSPLNT
27950	GRAFT-VERSUS-HOST NOS	99687	COMP INTESTINE TRANSPLNT
27951	AC GRAFT-VERSUS-HOST DIS	99688	COMP TP ORGAN-STEM CELL
27952	CHRONC GRAFT-VS-HOST DIS	99689	COMP OTH ORGAN TRANSPLNT
27953	AC ON CHRN GRFT-VS-HOST	V420	KIDNEY TRANSPLANT STATUS
2798	IMMUNE MECHANISM DIS NEC	V421	HEART TRANSPLANT STATUS
2799	IMMUNE MECHANISM DIS NOS	V426	LUNG TRANSPLANT STATUS
28409	CONST APLASTC ANEMIA NEC	V427	LIVER TRANSPLANT STATUS
2841	<i>PANCYTOPENIA</i>	V428	<i>OTHER SPECIFIED ORGAN OR TISSUE</i>
28411	ANTIN CHEMO INDCD PANCYT	V4281	TRANSPL STATUS-BNE MARROW
28412	OTH DRG INDCD PANCYTOPNA	V4282	TRSPL STS-PERIP STM CELL
28419	OTHER PANCYTOPENIA	V4283	TRANSPL STATUS-PANCREAS
2880	<i>AGRANULOCYTOSIS</i>	V4284	TRANSPL STATUS-INTESTINES
28800	NEUTROPENIA NOS	V4289	TRANSPL STATUS ORGAN NEC
28801	CONGENITAL NEUTROPENIA	V451	<i>RENAL DIALYSIS STATUS</i>
28802	CYCLIC NEUTROPENIA	V4511	RENAL DIALYSIS STATUS
28803	DRUG INDUCED NEUTROPENIA	V560	RENAL DIALYSIS ENCOUNTER
28809	NEUTROPENIA NEC	V561	FT/ADJ XTRCORP DIAL CATH
2881	FUNCTION DIS NEUTROPHILS	V562	FIT/ADJ PERIT DIAL CATH

<sup>1</sup> The procedure or diagnosis codes are continuously updated. The current list of ICD-9-CM codes is valid for October 2012 through September 2013. Italicized codes are not active in Fiscal Year 2013.

### ICD-9-CM Immunocompromised state procedure codes<sup>1</sup>:

0018	INFUS IMMUNOSUP ANTIBODY	336	COMB HEART/LUNG TRANSPLA
335	<i>LUNG TRANSPLANTATION</i>	375	<i>HEART TRANSPLANTATION</i>
3350	LUNG TRANSPLANT NOS	3751	HEART TRANSPLANTATION
3351	UNILAT LUNG TRANSPLANT	410	<i>OPERATIONS ON BONE MARROW AND SPLEEN</i>
3352	BILAT LUNG TRANSPLANT		

4100	BONE MARROW TRNSPLNT NOS	5051	AUXILIARY LIVER TRANSPL
4101	AUTO BONE MT W/O PURG	5059	LIVER TRANSPLANT NEC
4102	ALO BONE MARROW TRNSPLNT	5280	PANCREAT TRANSPLANT NOS
4103	ALLOGRFT BONE MARROW NOS	5281	REIMPLANT PANCREATIC TIS
4104	AUTO HEM STEM CT W/O PUR	5282	PANCREATIC HOMOTRANSPLAN
4105	ALLO HEM STEM CT W/O PUR	5283	PANCREATIC HETEROTRANSPL
4106	CORD BLD STEM CELL TRANS	5285	ALLOTRNSPLNT ISLETS LANG
4107	AUTO HEM STEM CT W PURG	5286	TRNSPLNT ISLETS LANG NOS
4108	ALLO HEM STEM CT W PURG	5569	KIDNEY TRANSPLANT NEC
4109	AUTO BONE MT W PURGING		

<sup>1</sup>The procedure or diagnosis codes are continuously updated. The current list of ICD-9-CM codes is valid for October 2012 through September 2013. Italicized codes are not active in Fiscal Year 2013.

## Appendix D – Definitions of Neonate, Newborn, Normal Newborn, and Outborn

A neonate is defined as any discharge with either:

- age in days at admission between zero and 28 days (inclusive); or
- age in days missing and age in years equal to zero and either:
  - an admission type of newborn (SID ATYPE=4); or
  - with any-listed ICD-9-CM diagnosis codes for in-hospital live birth; or
  - with any-listed ICD-9-CM diagnosis codes for neonatal observation and evaluation

### ICD-9-CM In-hospital live birth diagnosis codes:

V3000	SINGLE LB IN-HOSP W/O CS	V3401	OTH MULT LB-IN HOSP W CS
V3001	SINGLE LB IN-HOSP W CS	V3500	OTH MULT SB-HOSP W/O CS
V3100	TWIN-MATE LB-HOSP W/O CS	V3501	OTH MULT SB-IN HOSP W CS
V3101	TWIN-MATE LB-IN HOS W CS	V3600	MULT LB/SB-IN HOS W/O CS
V3200	TWIN-MATE SB-HOSP W/O CS	V3601	MULT LB/SB-IN HOSP W CS
V3201	TWIN-MATE SB-HOSP W CS	V3700	MULT BRTH NOS-HOS W/O CS
V3300	TWIN-NOS-IN HOSP W/O CS	V3701	MULT BIRTH NOS-HOSP W CS
V3301	TWIN-NOS-IN HOSP W CS	V3900	LIVEBORN NOS-HOSP W/O CS
V3400	OTH MULT LB-HOSP W/O CS	V3901	LIVEBORN NOS-HOSP W CS

### ICD-9-CM Neonatal observation and evaluation diagnosis codes:

V290	NB OBSRV SUSPCT INFECT	V293	NB OBS GENETC/METABL CND
V291	NB OBSRV SUSPCT NEURLGCL	V298	NB OBSRV OTH SUSPCT COND
V292	OBSRV NB SUSPC RESP COND	V299	NB OBSRV UNSP SUSPCT CND

A newborn is defined as any discharge meeting the definition of “neonate” (see above) with either:

- any-listed ICD-9-CM code for in-hospital live birth (see above) and age in days equal to zero or missing; or
- an admission type of newborn (SID ATYPE=4) and age in days equal to zero without any-listed ICD-9-CM diagnosis codes for out-of-hospital live birth; or
- an admission type of newborn (SID ATYPE=4) with point of origin for born inside this hospital (POINTOFORIGINUB04 code =5)

### ICD-9-CM Out-of-hospital live birth diagnosis codes:

V301	SINGL LIVEBRN-BEFORE ADM	V342	OTH MULTIPLE NB-NONHOSP
V302	SINGLE LIVEBORN-NONHOSP	V351	OTH MULT SB-BEFORE ADM
V311	TWIN, MATE LB-BEFORE ADM	V352	OTH MULTIPLE SB-NONHOSP
V312	TWIN, MATE LB-NONHOSP	V361	MULT NB/SB-BEFORE ADM
V321	TWIN, MATE SB-BEFORE ADM	V362	MULTIPLE NB/SB-NONHOSP
V322	TWIN, MATE SB-NONHOSP	V371	MULT BRTH NOS-BEFORE ADM
V331	TWIN NOS-BEFORE ADMISSN	V372	MULT BIRTH NOS-NONHOSP
V332	TWIN NOS-NONHOSP	V391	LIVEBORN NOS-BEFORE ADM
V341	OTH MULT NB-BEFORE ADM	V392	LIVEBORN NOS-NONHOSP

A normal newborn is defined as any discharge meeting the definition of “newborn” (see above) with a DRG code of 391 or a MS-DRG code 795.

An outborn is defined as any discharge meeting the definition of “neonate” (see above) that does not meet the definition of “newborn” (see above) with either:

- age in days less than 2 days and not missing; or
- an admission type of newborn (SID ATYPE=4) and age in days missing; or
- an admission type of newborn (SID ATYPE=4) and point of origin for born outside this hospital (POINTOFORIGINUB04 code =6)



# **PREVENTION QUALITY INDICATORS (PQI) PARAMETER ESTIMATES**

## **Version 4.5**

**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. HHSA290201200001C**

**Prepared by:**

**Battelle**  
505 King Avenue  
Columbus, OH 43201

**May 2013**

## Table of Contents

Executive Summary .....	ii
-------------------------	----

### List of Risk Adjustment Coefficient Tables

Table 1. PQI #1 Diabetes Short-Term Complications Admission Rate .....	1
Table 2. PQI #2 Perforated Appendix Admission Rate .....	3
Table 3. PQI #3 Diabetes Long-Term Complications Admission Rate.....	5
Table 4. PQI #5 Chronic Obstructive Pulmonary Disease (COPD) or Asthma in Older Adults Admission Rate .....	7
Table 5. PQI #7 Hypertension Admission Rate.....	8
Table 6. PQI #8 Heart Failure Admission Rate .....	10
Table 7. PQI #10 Dehydration Admission Rate .....	12
Table 8. PQI #11 Bacterial Pneumonia Admission Rate .....	14
Table 9. PQI #12 Urinary Tract Infection Admission Rate .....	16
Table 10. PQI #13 Angina Without Procedure Admission Rate .....	18
Table 11. PQI #14 Uncontrolled Diabetes Admission Rate .....	20
Table 12. PQI #15 Asthma in Younger Adults Admission Rate .....	22
Table 13. PQI #16 Lower-Extremity Amputation among Patients with Diabetes Rate .....	23
Table 14. PQI #90 Prevention Quality Overall Composite .....	25
Table 15. PQI #91 - Prevention Quality Acute Composite .....	27
Table 16. PQI #92 - Prevention Quality Chronic Composite .....	29

### List of Tables Describing Categorical Covariates

Table A.1. Population Age Categories .....	31
--	----

## Executive Summary

This document provides the risk adjustment covariates and coefficients for relevant Agency for Healthcare Research and Quality (AHRQ) Quality Indicators™ (QI) Prevention Quality Indicators (PQI). The parameter estimates derived for the AHRQ QI are based on analysis of the 2010 Agency for Healthcare Research and Quality's Healthcare Cost and Utilization Project (HCUP) State Inpatient Databases (SID). HCUP is a family of health care databases and related software tools and products developed through a Federal-State-Industry partnership<sup>1</sup>. HCUP includes the largest collection of longitudinal hospital care data in the United States, with all-payer, encounter-level information beginning in 1988. The SID contain all-payer, encounter-level information on inpatient discharges, including clinical and resource information typically found on a billing record, such as patient demographics, up to 30 *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)* diagnoses and procedures, length of stay (LOS), expected payer, admission and discharge dates and discharge disposition. In 2010, the HCUP databases represent more than 95 percent of all annual discharges in the U.S.<sup>2</sup>

These coefficients are used by the prediction module to calculate risk-adjusted rates that account for differences in patient populations across areas. Covariates that are considered as potential risk adjusters include gender and age and the interaction of gender and age. Descriptions of the population age categories are provided in the Table A.1. Every covariate in every model is a binary indicator variable, coded using 0 or 1. The AHRQ QI software user does not need to manipulate or adjust these coefficients; rather this document is intended to make it transparent to the user how the risk adjusted QI rates are calculated.

---

<sup>1</sup> The AHRQ QI program would like to acknowledge the HCUP Partner organizations that participated in the HCUP SID: **Alaska** State Hospital and Nursing Home Association, **Arizona** Department of Health Services, **Arkansas** Department of Health, **California** Office of Statewide Health Planning and Development, **Colorado** Hospital Association, **Connecticut** Hospital Association, **Florida** Agency for Health Care Administration, **Georgia** Hospital Association, **Hawaii** Health Information Corporation, **Illinois** Department of Public Health, **Indiana** Hospital Association, **Iowa** Hospital Association, **Kansas** Hospital Association, **Kentucky** Cabinet for Health and Family Services, **Louisiana** Department of Health and Hospitals, **Maine** Health Data Organization, **Maryland** Health Services Cost Review Commission, **Massachusetts** Center for Health Information and Analysis, **Michigan** Health & Hospital Association, **Minnesota** Hospital Association (provides data for Minnesota and North Dakota), **Mississippi** Department of Health, **Missouri** Hospital Industry Data Institute, **Montana** MHA - An Association of Montana Health Care Providers, **Nebraska** Hospital Association, **Nevada** Department of Health and Human Services, **New Hampshire** Department of Health & Human Services, **New Jersey** Department of Health, **New Mexico** Department of Health, **New York** State Department of Health, **North Carolina** Department of Health and Human Services, **North Dakota** (data provided by the Minnesota Hospital Association), **Ohio** Hospital Association, **Oklahoma** State Department of Health, **Oregon** Association of Hospitals and Health Systems, **Oregon** Health Policy and Research, **Pennsylvania** Health Care Cost Containment Council, **Rhode Island** Department of Health, **South Carolina** Budget & Control Board, **South Dakota** Association of Healthcare Organizations, **Tennessee** Hospital Association, **Texas** Department of State Health Services, **Utah** Department of Health, **Vermont** Association of Hospitals and Health Systems, **Virginia** Health Information, **Washington** State Department of Health, **West Virginia** Health Care Authority, **Wisconsin** Department of Health Services, **Wyoming** Hospital Association

<sup>2</sup> The states included in the analysis are Alaska, Arkansas, Arizona, California, Colorado, Connecticut, Florida, Georgia, Hawaii, Iowa, Illinois, Indiana, Kansas, Kentucky, Louisiana, Massachusetts, Maryland, Maine, Michigan, Minnesota, Missouri, Montana, North Carolina, Nebraska, New Jersey, New Mexico, Nevada, New York, Ohio, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Carolina, South Dakota, Tennessee, Texas, Utah, Virginia, Vermont, Washington, Wisconsin, West Virginia, and Wyoming.

Additional information on the risk adjustment process can be found in *Quality Indicator Empirical Methods*, available on the AHRQ QI™ website.  
(<http://www.qualityindicators.ahrq.gov/modules/Default.aspx>)



**Table 1. Risk Adjustment Coefficients for PQI #1 Diabetes Short-Term Complications Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-8.0120	0.0416	37143.58	< 0.0001
SEX	Female	1	-0.1430	0.0519	7.65	0.0057
AGE5	Male, Age 18-24	1	0.9129	0.0425	460.50	< 0.0001
AGE6	Male, Age 25-29	1	0.8191	0.0431	361.07	< 0.0001
AGE7	Male, Age 30-34	1	0.8126	0.0432	353.71	< 0.0001
AGE8	Male, Age 35-39	1	0.7570	0.0433	305.53	< 0.0001
AGE9	Male, Age 40-44	1	0.8595	0.0431	398.02	< 0.0001
AGE10	Male, Age 45-49	1	0.8403	0.0430	381.81	< 0.0001
AGE11	Male, Age 50-54	1	0.7143	0.0432	273.06	< 0.0001
AGE12	Male, Age 55-59	1	0.5363	0.0438	149.78	< 0.0001
AGE13	Male, Age 60-64	1	0.3375	0.0448	56.84	< 0.0001
AGE14	Male, Age 65-69	1	0.2139	0.0465	21.18	< 0.0001
AGE15	Male, Age 70-74	1	0.1144	0.0488	5.48	0.0192
AGE16	Male, Age 75-79	1	0.1127	0.0511	4.86	0.0274
AGE17	Male, Age 80-84	1	0.0655	0.0548	1.43	0.2323
AGE5	Female, Age 18-24	1	0.3665	0.0534	47.17	< 0.0001
AGE6	Female, Age 25-29	1	0.0813	0.0545	2.23	0.1358
AGE7	Female, Age 30-34	1	-0.1260	0.0549	5.29	0.0215
AGE8	Female, Age 35-39	1	-0.0210	0.0549	0.16	0.6935
AGE9	Female, Age 40-44	1	-0.1800	0.0548	10.90	0.0010
AGE10	Female, Age 45-49	1	-0.1600	0.0546	8.67	0.0032
AGE11	Female, Age 50-54	1	-0.0630	0.0548	1.35	0.2454
AGE12	Female, Age 55-59	1	0.0553	0.0556	0.99	0.3192
AGE13	Female, Age 60-64	1	0.0926	0.0569	2.65	0.1036
AGE14	Female, Age 65-69	1	0.0852	0.0595	2.06	0.1517

(CONTINUED)

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	0.1170	0.0627	3.49	0.0619
AGE16	Female, Age 75-79	1	0.1120	0.0654	2.93	0.0870
AGE17	Female, Age 80-84	1	0.1198	0.0696	2.96	0.0852

c-statistic: Measures of association between the observed and predicted values were not calculated because the predicted probabilities are indistinguishable when they are classified into intervals of length 0.002.

**Table 2. Risk Adjustment Coefficients for PQI #2 Perforated Appendix Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	0.2954	0.0620	22.71	< 0.0001
SEX	Female	1	0.0065	0.0808	0.01	0.9357
AGE5	Male, Age 18-24	1	-1.6940	0.0643	694.33	< 0.0001
AGE6	Male, Age 25-29	1	-1.6690	0.0659	642.62	< 0.0001
AGE7	Male, Age 30-34	1	-1.5780	0.0661	571.23	< 0.0001
AGE8	Male, Age 35-39	1	-1.4090	0.0662	453.20	< 0.0001
AGE9	Male, Age 40-44	1	-1.1410	0.0660	298.69	< 0.0001
AGE10	Male, Age 45-49	1	-0.8710	0.0657	175.99	< 0.0001
AGE11	Male, Age 50-54	1	-0.6060	0.0660	84.50	< 0.0001
AGE12	Male, Age 55-59	1	-0.4430	0.0666	44.36	< 0.0001
AGE13	Male, Age 60-64	1	-0.2830	0.0675	17.64	< 0.0001
AGE14	Male, Age 65-69	1	-0.2930	0.0690	18.09	< 0.0001
AGE15	Male, Age 70-74	1	-0.1900	0.0718	7.04	0.0080
AGE16	Male, Age 75-79	1	-0.0760	0.0748	1.04	0.3078
AGE17	Male, Age 80-84	1	-0.0280	0.0805	0.13	0.7207
AGE5	Female, Age 18-24	1	-0.4090	0.0856	22.88	< 0.0001
AGE6	Female, Age 25-29	1	-0.3510	0.0884	15.75	< 0.0001
AGE7	Female, Age 30-34	1	-0.3100	0.0887	12.28	0.0005
AGE8	Female, Age 35-39	1	-0.2950	0.0886	11.11	0.0009
AGE9	Female, Age 40-44	1	-0.3340	0.0879	14.47	0.0001
AGE10	Female, Age 45-49	1	-0.3160	0.0869	13.23	0.0003
AGE11	Female, Age 50-54	1	-0.2820	0.0869	10.56	0.0012
AGE12	Female, Age 55-59	1	-0.3360	0.0879	14.65	0.0001
AGE13	Female, Age 60-64	1	-0.3090	0.0892	12.07	0.0005
AGE14	Female, Age 65-69	1	-0.1520	0.0916	2.78	0.0952

(CONTINUED)

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	-0.1640	0.0954	2.98	0.0841
AGE16	Female, Age 75-79	1	-0.1960	0.1005	3.83	0.0503
AGE17	Female, Age 80-84	1	-0.0600	0.1073	0.32	0.5717

c-statistic = 0.671

**Table 3. Risk Adjustment Coefficients for PQI #3 Diabetes Long-Term Complications Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-5.5950	0.0125	201807.6	< 0.0001
SEX	Female	1	-0.3420	0.0162	448.74	< 0.0001
AGE5	Male, Age 18-24	1	-3.8890	0.0322	14621.39	< 0.0001
AGE6	Male, Age 25-29	1	-2.7110	0.0234	13380.76	< 0.0001
AGE7	Male, Age 30-34	1	-2.1790	0.0200	11849.91	< 0.0001
AGE8	Male, Age 35-39	1	-1.7800	0.0179	9878.12	< 0.0001
AGE9	Male, Age 40-44	1	-1.4260	0.0163	7624.07	< 0.0001
AGE10	Male, Age 45-49	1	-1.0530	0.0151	4893.88	< 0.0001
AGE11	Male, Age 50-54	1	-0.8220	0.0146	3171.17	< 0.0001
AGE12	Male, Age 55-59	1	-0.6210	0.0145	1840.11	< 0.0001
AGE13	Male, Age 60-64	1	-0.5290	0.0146	1312.98	< 0.0001
AGE14	Male, Age 65-69	1	-0.3510	0.0149	554.71	< 0.0001
AGE15	Male, Age 70-74	1	-0.1770	0.0153	135.04	< 0.0001
AGE16	Male, Age 75-79	1	-0.0800	0.0158	25.56	< 0.0001
AGE17	Male, Age 80-84	1	0.0114	0.0166	0.47	0.4936
AGE5	Female, Age 18-24	1	0.9414	0.0406	538.20	< 0.0001
AGE6	Female, Age 25-29	1	0.6806	0.0307	491.64	< 0.0001
AGE7	Female, Age 30-34	1	0.3276	0.0275	141.74	< 0.0001
AGE8	Female, Age 35-39	1	0.2244	0.0247	82.24	< 0.0001
AGE9	Female, Age 40-44	1	0.0539	0.0228	5.58	0.0182
AGE10	Female, Age 45-49	1	-0.0550	0.0209	7.03	0.0080
AGE11	Female, Age 50-54	1	-0.1090	0.0202	29.28	< 0.0001
AGE12	Female, Age 55-59	1	-0.1050	0.0199	28.29	< 0.0001
AGE13	Female, Age 60-64	1	-0.0220	0.0199	1.29	0.2565
AGE14	Female, Age 65-69	1	-0.0130	0.0204	0.41	0.5202

(CONTINUED)

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	0.0066	0.0208	0.10	0.7504
AGE16	Female, Age 75-79	1	0.1087	0.0212	26.26	< 0.0001
AGE17	Female, Age 80-84	1	0.0678	0.0221	9.43	0.0021

c-statistic = 0.621

**Table 4. Risk Adjustment Coefficients for PQI #5 Chronic Obstructive Pulmonary Disease (COPD) or Asthma in Older Adults Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-4.3460	0.0067	416764.7	< 0.0001
SEX	Female	1	-0.1860	0.0084	485.31	< 0.0001
AGE9	Male, Age 40-44	1	-2.7240	0.0128	45581.78	< 0.0001
AGE10	Male, Age 45-49	1	-2.2290	0.0106	44342.11	< 0.0001
AGE11	Male, Age 50-54	1	-1.7620	0.0093	35256.24	< 0.0001
AGE12	Male, Age 55-59	1	-1.4450	0.0090	25777.49	< 0.0001
AGE13	Male, Age 60-64	1	-1.1350	0.0087	16905.55	< 0.0001
AGE14	Male, Age 65-69	1	-0.6480	0.0084	5863.57	< 0.0001
AGE15	Male, Age 70-74	1	-0.3400	0.0085	1606.06	< 0.0001
AGE16	Male, Age 75-79	1	-0.1340	0.0086	242.34	< 0.0001
AGE17	Male, Age 80-84	1	0.0080	0.0089	0.81	0.3680
AGE9	Female, Age 40-44	1	0.9917	0.0155	4071.22	< 0.0001
AGE10	Female, Age 45-49	1	0.9108	0.0130	4875.29	< 0.0001
AGE11	Female, Age 50-54	1	0.7630	0.0117	4225.22	< 0.0001
AGE12	Female, Age 55-59	1	0.5943	0.0114	2717.51	< 0.0001
AGE13	Female, Age 60-64	1	0.5180	0.0111	2177.72	< 0.0001
AGE14	Female, Age 65-69	1	0.4002	0.0108	1365.78	< 0.0001
AGE15	Female, Age 70-74	1	0.3130	0.0109	824.02	< 0.0001
AGE16	Female, Age 75-79	1	0.2444	0.0111	486.25	< 0.0001
AGE17	Female, Age 80-84	1	0.1131	0.0115	96.37	< 0.0001

c-statistic = 0.689

**Table 5. Risk Adjustment Coefficients for PQI #7 Hypertension Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-6.5220	0.0198	108971.0	< 0.0001
SEX	Female	1	0.7482	0.0219	1164.89	< 0.0001
AGE5	Male, Age 18-24	1	-3.9890	0.0534	5589.52	< 0.0001
AGE6	Male, Age 25-29	1	-2.7450	0.0377	5302.76	< 0.0001
AGE7	Male, Age 30-34	1	-2.0620	0.0307	4512.30	< 0.0001
AGE8	Male, Age 35-39	1	-1.4190	0.0261	2952.60	< 0.0001
AGE9	Male, Age 40-44	1	-1.0320	0.0241	1835.58	< 0.0001
AGE10	Male, Age 45-49	1	-0.8870	0.0233	1449.42	< 0.0001
AGE11	Male, Age 50-54	1	-0.7550	0.0230	1082.74	< 0.0001
AGE12	Male, Age 55-59	1	-0.7490	0.0234	1027.44	< 0.0001
AGE13	Male, Age 60-64	1	-0.7540	0.0240	989.59	< 0.0001
AGE14	Male, Age 65-69	1	-0.6430	0.0249	668.03	< 0.0001
AGE15	Male, Age 70-74	1	-0.4790	0.0257	348.98	< 0.0001
AGE16	Male, Age 75-79	1	-0.3690	0.0267	191.87	< 0.0001
AGE17	Male, Age 80-84	1	-0.2050	0.0277	54.88	< 0.0001
AGE5	Female, Age 18-24	1	-0.5520	0.0711	60.34	< 0.0001
AGE6	Female, Age 25-29	1	-0.9180	0.0525	305.79	< 0.0001
AGE7	Female, Age 30-34	1	-0.8040	0.0402	399.59	< 0.0001
AGE8	Female, Age 35-39	1	-0.9350	0.0335	780.32	< 0.0001
AGE9	Female, Age 40-44	1	-0.9210	0.0299	947.24	< 0.0001
AGE10	Female, Age 45-49	1	-0.7860	0.0281	784.14	< 0.0001
AGE11	Female, Age 50-54	1	-0.7560	0.0274	762.01	< 0.0001
AGE12	Female, Age 55-59	1	-0.7050	0.0279	640.16	< 0.0001
AGE13	Female, Age 60-64	1	-0.5660	0.0284	397.43	< 0.0001
AGE14	Female, Age 65-69	1	-0.4130	0.0292	199.81	< 0.0001

(CONTINUED)



**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	-0.2500	0.0298	70.52	< 0.0001
AGE16	Female, Age 75-79	1	-0.0800	0.0305	7.04	0.0080
AGE17	Female, Age 80-84	1	-0.0030	0.0313	0.01	0.9102

c-statistic = 0.558

**Table 6. Risk Adjustment Coefficients for PQI #8 Heart Failure Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-3.1540	0.0038	682427.4	< 0.0001
SEX	Female	1	-0.1900	0.0048	1578.70	< 0.0001
AGE5	Male, Age 18-24	1	-6.7190	0.0362	34395.85	< 0.0001
AGE6	Male, Age 25-29	1	-5.6980	0.0264	46672.32	< 0.0001
AGE7	Male, Age 30-34	1	-5.0260	0.0196	65908.47	< 0.0001
AGE8	Male, Age 35-39	1	-4.4470	0.0149	88849.60	< 0.0001
AGE9	Male, Age 40-44	1	-3.9070	0.0114	116521.9	< 0.0001
AGE10	Male, Age 45-49	1	-3.4040	0.0089	144496.9	< 0.0001
AGE11	Male, Age 50-54	1	-2.9660	0.0076	152042.4	< 0.0001
AGE12	Male, Age 55-59	1	-2.5960	0.0069	137831.4	< 0.0001
AGE13	Male, Age 60-64	1	-2.2370	0.0065	116869.1	< 0.0001
AGE14	Male, Age 65-69	1	-1.7860	0.0063	80485.94	< 0.0001
AGE15	Male, Age 70-74	1	-1.3760	0.0061	50199.35	< 0.0001
AGE16	Male, Age 75-79	1	-0.9520	0.0059	25847.68	< 0.0001
AGE17	Male, Age 80-84	1	-0.5220	0.0057	8186.83	< 0.0001
AGE5	Female, Age 18-24	1	0.0043	0.0544	0.01	0.9360
AGE6	Female, Age 25-29	1	-0.1780	0.0413	18.61	< 0.0001
AGE7	Female, Age 30-34	1	-0.2970	0.0315	88.70	< 0.0001
AGE8	Female, Age 35-39	1	-0.3180	0.0240	176.71	< 0.0001
AGE9	Female, Age 40-44	1	-0.3340	0.0183	334.97	< 0.0001
AGE10	Female, Age 45-49	1	-0.2700	0.0138	383.98	< 0.0001
AGE11	Female, Age 50-54	1	-0.3020	0.0116	680.17	< 0.0001
AGE12	Female, Age 55-59	1	-0.2590	0.0104	625.03	< 0.0001
AGE13	Female, Age 60-64	1	-0.1550	0.0093	273.39	< 0.0001
AGE14	Female, Age 65-69	1	-0.1300	0.0088	216.03	< 0.0001

(CONTINUED)

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	-0.0690	0.0084	66.93	< 0.0001
AGE16	Female, Age 75-79	1	-0.0520	0.0079	42.85	< 0.0001
AGE17	Female, Age 80-84	1	-0.0430	0.0075	32.84	< 0.0001

c-statistic = 0.853

**Table 7. Risk Adjustment Coefficients for PQI #10 Dehydration Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-4.6880	0.0079	347064.2	< 0.0001
SEX	Female	1	0.0309	0.0096	10.24	0.0014
AGE5	Male, Age 18-24	1	-3.6190	0.0183	39157.76	< 0.0001
AGE6	Male, Age 25-29	1	-3.4880	0.0202	29699.20	< 0.0001
AGE7	Male, Age 30-34	1	-3.4080	0.0201	28870.56	< 0.0001
AGE8	Male, Age 35-39	1	-3.2620	0.0189	29728.11	< 0.0001
AGE9	Male, Age 40-44	1	-3.0650	0.0172	31775.15	< 0.0001
AGE10	Male, Age 45-49	1	-2.7740	0.0150	34199.01	< 0.0001
AGE11	Male, Age 50-54	1	-2.4910	0.0137	33059.47	< 0.0001
AGE12	Male, Age 55-59	1	-2.2310	0.0132	28727.89	< 0.0001
AGE13	Male, Age 60-64	1	-1.9870	0.0128	23989.69	< 0.0001
AGE14	Male, Age 65-69	1	-1.6200	0.0127	16333.08	< 0.0001
AGE15	Male, Age 70-74	1	-1.2800	0.0126	10300.12	< 0.0001
AGE16	Male, Age 75-79	1	-0.9050	0.0123	5404.61	< 0.0001
AGE17	Male, Age 80-84	1	-0.5240	0.0121	1869.06	< 0.0001
AGE5	Female, Age 18-24	1	0.1514	0.0245	38.15	< 0.0001
AGE6	Female, Age 25-29	1	0.1473	0.0271	29.48	< 0.0001
AGE7	Female, Age 30-34	1	0.2257	0.0264	73.21	< 0.0001
AGE8	Female, Age 35-39	1	0.2560	0.0246	107.91	< 0.0001
AGE9	Female, Age 40-44	1	0.2575	0.0223	133.15	< 0.0001
AGE10	Female, Age 45-49	1	0.2162	0.0194	123.61	< 0.0001
AGE11	Female, Age 50-54	1	0.1582	0.0178	78.94	< 0.0001
AGE12	Female, Age 55-59	1	0.1354	0.0170	63.05	< 0.0001
AGE13	Female, Age 60-64	1	0.1525	0.0165	85.38	< 0.0001
AGE14	Female, Age 65-69	1	0.1201	0.0163	54.39	< 0.0001

(CONTINUED)

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	0.1311	0.0161	66.62	< 0.0001
AGE16	Female, Age 75-79	1	0.1137	0.0155	53.72	< 0.0001
AGE17	Female, Age 80-84	1	0.0944	0.0150	39.39	< 0.0001

c-statistic = 0.709

**Table 8. Risk Adjustment Coefficients for PQI #11 Bacterial Pneumonia Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-3.4440	0.0043	621558.0	< 0.0001
SEX	Female	1	-0.3160	0.0056	3174.15	< 0.0001
AGE5	Male, Age 18-24	1	-4.5270	0.0146	96219.66	< 0.0001
AGE6	Male, Age 25-29	1	-4.2860	0.0155	76235.66	< 0.0001
AGE7	Male, Age 30-34	1	-4.0210	0.0141	80999.04	< 0.0001
AGE8	Male, Age 35-39	1	-3.7810	0.0127	88266.17	< 0.0001
AGE9	Male, Age 40-44	1	-3.4920	0.0110	100049.1	< 0.0001
AGE10	Male, Age 45-49	1	-3.1860	0.0094	113304.5	< 0.0001
AGE11	Male, Age 50-54	1	-2.8610	0.0084	115090.4	< 0.0001
AGE12	Male, Age 55-59	1	-2.6010	0.0080	103921.2	< 0.0001
AGE13	Male, Age 60-64	1	-2.2810	0.0076	89061.40	< 0.0001
AGE14	Male, Age 65-69	1	-1.8000	0.0072	61238.57	< 0.0001
AGE15	Male, Age 70-74	1	-1.3560	0.0070	37359.80	< 0.0001
AGE16	Male, Age 75-79	1	-0.9160	0.0067	18525.42	< 0.0001
AGE17	Male, Age 80-84	1	-0.5110	0.0066	5998.18	< 0.0001
AGE5	Female, Age 18-24	1	0.3590	0.0205	306.70	< 0.0001
AGE6	Female, Age 25-29	1	0.4537	0.0212	456.37	< 0.0001
AGE7	Female, Age 30-34	1	0.4559	0.0192	561.65	< 0.0001
AGE8	Female, Age 35-39	1	0.5022	0.0171	862.44	< 0.0001
AGE9	Female, Age 40-44	1	0.4900	0.0148	1090.16	< 0.0001
AGE10	Female, Age 45-49	1	0.4856	0.0127	1470.33	< 0.0001
AGE11	Female, Age 50-54	1	0.4317	0.0113	1451.56	< 0.0001
AGE12	Female, Age 55-59	1	0.3646	0.0109	1117.45	< 0.0001
AGE13	Female, Age 60-64	1	0.3373	0.0103	1066.76	< 0.0001
AGE14	Female, Age 65-69	1	0.2476	0.0098	627.02	< 0.0001

(CONTINUED)

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	0.1627	0.0095	289.31	< 0.0001
AGE16	Female, Age 75-79	1	0.0650	0.0091	50.41	< 0.0001
AGE17	Female, Age 80-84	1	0.0220	0.0088	6.21	0.0127

c-statistic = 0.791

**Table 9. Risk Adjustment Coefficients for PQI #12 Urinary Tract Infection Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-4.2000	0.0062	448796.2	< 0.0001
SEX	Female	1	0.4861	0.0071	4608.63	< 0.0001
AGE5	Male, Age 18-24	1	-4.9810	0.0262	36038.30	< 0.0001
AGE6	Male, Age 25-29	1	-4.5870	0.0260	31084.11	< 0.0001
AGE7	Male, Age 30-34	1	-4.4270	0.0248	31852.32	< 0.0001
AGE8	Male, Age 35-39	1	-4.2420	0.0228	34546.98	< 0.0001
AGE9	Male, Age 40-44	1	-3.9970	0.0200	39820.01	< 0.0001
AGE10	Male, Age 45-49	1	-3.7020	0.0170	47241.78	< 0.0001
AGE11	Male, Age 50-54	1	-3.4280	0.0153	50220.94	< 0.0001
AGE12	Male, Age 55-59	1	-3.0860	0.0141	48150.76	< 0.0001
AGE13	Male, Age 60-64	1	-2.6900	0.0128	43926.18	< 0.0001
AGE14	Male, Age 65-69	1	-2.1390	0.0118	32767.73	< 0.0001
AGE15	Male, Age 70-74	1	-1.6010	0.0110	21287.74	< 0.0001
AGE16	Male, Age 75-79	1	-1.1000	0.0103	11489.28	< 0.0001
AGE17	Male, Age 80-84	1	-0.5830	0.0097	3615.56	< 0.0001
AGE5	Female, Age 18-24	1	1.8945	0.0276	4701.51	< 0.0001
AGE6	Female, Age 25-29	1	1.4139	0.0280	2541.52	< 0.0001
AGE7	Female, Age 30-34	1	1.1708	0.0272	1857.62	< 0.0001
AGE8	Female, Age 35-39	1	0.9942	0.0253	1539.76	< 0.0001
AGE9	Female, Age 40-44	1	0.7746	0.0227	1164.90	< 0.0001
AGE10	Female, Age 45-49	1	0.5099	0.0198	662.09	< 0.0001
AGE11	Female, Age 50-54	1	0.3425	0.0181	357.51	< 0.0001
AGE12	Female, Age 55-59	1	0.1811	0.0169	114.55	< 0.0001
AGE13	Female, Age 60-64	1	0.1076	0.0155	48.04	< 0.0001
AGE14	Female, Age 65-69	1	0.0119	0.0144	0.68	0.4079

(CONTINUED)



**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	0.0127	0.0132	0.92	0.3378
AGE16	Female, Age 75-79	1	0.0303	0.0122	6.22	0.0127
AGE17	Female, Age 80-84	1	0.0212	0.0113	3.49	0.0616

c-statistic = 0.771

**Table 10. Risk Adjustment Coefficients for PQI #13 Angina Without Procedure Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-7.3400	0.0297	61012.26	< 0.0001
SEX	Female	1	-0.0020	0.0362	0.01	0.9403
AGE5	Male, Age 18-24	1	-5.0690	0.1314	1487.55	< 0.0001
AGE6	Male, Age 25-29	1	-4.2300	0.1058	1598.82	< 0.0001
AGE7	Male, Age 30-34	1	-3.1220	0.0670	2169.64	< 0.0001
AGE8	Male, Age 35-39	1	-2.2060	0.0483	2084.06	< 0.0001
AGE9	Male, Age 40-44	1	-1.4990	0.0396	1431.30	< 0.0001
AGE10	Male, Age 45-49	1	-1.0960	0.0362	917.12	< 0.0001
AGE11	Male, Age 50-54	1	-0.8730	0.0351	618.45	< 0.0001
AGE12	Male, Age 55-59	1	-0.7420	0.0351	446.36	< 0.0001
AGE13	Male, Age 60-64	1	-0.6820	0.0357	366.06	< 0.0001
AGE14	Male, Age 65-69	1	-0.5680	0.0369	237.17	< 0.0001
AGE15	Male, Age 70-74	1	-0.4920	0.0387	162.03	< 0.0001
AGE16	Male, Age 75-79	1	-0.3350	0.0398	70.85	< 0.0001
AGE17	Male, Age 80-84	1	-0.1960	0.0416	22.31	< 0.0001
AGE5	Female, Age 18-24	1	-0.3480	0.2052	2.89	0.0892
AGE6	Female, Age 25-29	1	-0.5090	0.1708	8.89	0.0029
AGE7	Female, Age 30-34	1	-0.2320	0.0975	5.69	0.0170
AGE8	Female, Age 35-39	1	-0.4300	0.0706	37.07	< 0.0001
AGE9	Female, Age 40-44	1	-0.2600	0.0538	23.47	< 0.0001
AGE10	Female, Age 45-49	1	-0.2320	0.0477	23.83	< 0.0001
AGE11	Female, Age 50-54	1	-0.1220	0.0452	7.37	0.0066
AGE12	Female, Age 55-59	1	-0.1610	0.0453	12.65	0.0004
AGE13	Female, Age 60-64	1	-0.0900	0.0458	3.92	0.0476
AGE14	Female, Age 65-69	1	-0.0530	0.0474	1.26	0.2621

(CONTINUED)

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	0.1281	0.0489	6.87	0.0088
AGE16	Female, Age 75-79	1	0.0725	0.0503	2.08	0.1492
AGE17	Female, Age 80-84	1	0.0605	0.0519	1.36	0.2434

c-statistic: Measures of association between the observed and predicted values were not calculated because the predicted probabilities are indistinguishable when they are classified into intervals of length 0.002.

**Table 11. Risk Adjustment Coefficients for PQI #14 Uncontrolled Diabetes Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-7.9190	0.0397	39816.12	< 0.0001
SEX	Female	1	-0.0760	0.0490	2.43	0.1190
AGE5	Male, Age 18-24	1	-2.0100	0.0543	1370.92	< 0.0001
AGE6	Male, Age 25-29	1	-1.5660	0.0535	858.47	< 0.0001
AGE7	Male, Age 30-34	1	-1.2030	0.0502	574.69	< 0.0001
AGE8	Male, Age 35-39	1	-0.9080	0.0478	361.60	< 0.0001
AGE9	Male, Age 40-44	1	-0.5930	0.0455	170.19	< 0.0001
AGE10	Male, Age 45-49	1	-0.4370	0.0444	97.13	< 0.0001
AGE11	Male, Age 50-54	1	-0.3350	0.0440	57.91	< 0.0001
AGE12	Male, Age 55-59	1	-0.3250	0.0446	53.17	< 0.0001
AGE13	Male, Age 60-64	1	-0.3950	0.0458	74.60	< 0.0001
AGE14	Male, Age 65-69	1	-0.2740	0.0471	33.94	< 0.0001
AGE15	Male, Age 70-74	1	-0.1060	0.0482	4.89	0.0269
AGE16	Male, Age 75-79	1	-0.0670	0.0504	1.80	0.1791
AGE17	Male, Age 80-84	1	0.0743	0.0522	2.02	0.1548
AGE5	Female, Age 18-24	1	0.1494	0.0715	4.37	0.0367
AGE6	Female, Age 25-29	1	-0.0510	0.0719	0.51	0.4748
AGE7	Female, Age 30-34	1	-0.2460	0.0683	13.04	0.0003
AGE8	Female, Age 35-39	1	-0.2360	0.0638	13.78	0.0002
AGE9	Female, Age 40-44	1	-0.2770	0.0600	21.43	< 0.0001
AGE10	Female, Age 45-49	1	-0.1930	0.0575	11.35	0.0008
AGE11	Female, Age 50-54	1	-0.0200	0.0561	0.13	0.7145
AGE12	Female, Age 55-59	1	0.0068	0.0568	0.01	0.9043
AGE13	Female, Age 60-64	1	0.1181	0.0582	4.12	0.0423
AGE14	Female, Age 65-69	1	0.1301	0.0598	4.73	0.0297

(CONTINUED)

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	0.1419	0.0612	5.39	0.0203
AGE16	Female, Age 75-79	1	0.1593	0.0636	6.27	0.0123
AGE17	Female, Age 80-84	1	0.1214	0.0655	3.43	0.0638

c-statistic: Measures of association between the observed and predicted values were not calculated because the predicted probabilities are indistinguishable when they are classified into intervals of length 0.002.

**Table 12. Risk Adjustment Coefficients for PQI #15 Asthma in Younger Adults Admission Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-7.7670	0.0157	245895.2	< 0.0001
SEX	Female	1	0.9582	0.0184	2709.77	< 0.0001
AGE5	Male, Age 18-24	1	-0.4420	0.0222	398.22	< 0.0001
AGE6	Male, Age 25-29	1	-0.3380	0.0238	201.90	< 0.0001
AGE7	Male, Age 30-34	1	-0.1820	0.0232	62.01	< 0.0001
AGE5	Female, Age 18-24	1	-0.3900	0.0270	208.82	< 0.0001
AGE6	Female, Age 25-29	1	-0.2310	0.0287	65.09	< 0.0001
AGE7	Female, Age 30-34	1	-0.1040	0.0275	14.28	0.0002

c-statistic: Measures of association between the observed and predicted values were not calculated because the predicted probabilities are indistinguishable when they are classified into intervals of length 0.002.

**Table 13. Risk Adjustment Coefficients for PQI #16 Lower-Extremity Amputation among Patients with Diabetes Rate**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-7.3660	0.0301	59872.68	< 0.0001
SEX	Female	1	-0.6190	0.0415	222.24	< 0.0001
AGE5	Male, Age 18-24	1	-6.8480	0.3172	466.21	< 0.0001
AGE6	Male, Age 25-29	1	-4.5300	0.1233	1351.06	< 0.0001
AGE7	Male, Age 30-34	1	-3.5490	0.0812	1912.81	< 0.0001
AGE8	Male, Age 35-39	1	-2.7030	0.0579	2176.78	< 0.0001
AGE9	Male, Age 40-44	1	-2.0680	0.0464	1986.23	< 0.0001
AGE10	Male, Age 45-49	1	-1.3900	0.0387	1292.39	< 0.0001
AGE11	Male, Age 50-54	1	-0.9370	0.0359	681.85	< 0.0001
AGE12	Male, Age 55-59	1	-0.5580	0.0347	258.48	< 0.0001
AGE13	Male, Age 60-64	1	-0.3580	0.0346	107.25	< 0.0001
AGE14	Male, Age 65-69	1	-0.1050	0.0349	9.09	0.0026
AGE15	Male, Age 70-74	1	0.0465	0.0357	1.70	0.1927
AGE16	Male, Age 75-79	1	0.0715	0.0373	3.69	0.0549
AGE17	Male, Age 80-84	1	0.0771	0.0396	3.79	0.0514
AGE5	Female, Age 18-24	1	0.9274	0.4220	4.83	0.0280
AGE6	Female, Age 25-29	1	0.1970	0.1955	1.02	0.3136
AGE7	Female, Age 30-34	1	0.2567	0.1249	4.22	0.0399
AGE8	Female, Age 35-39	1	0.0112	0.0929	0.01	0.9043
AGE9	Female, Age 40-44	1	-0.0130	0.0728	0.03	0.8524
AGE10	Female, Age 45-49	1	-0.1990	0.0601	11.02	0.0009
AGE11	Female, Age 50-54	1	-0.2160	0.0544	15.90	< 0.0001
AGE12	Female, Age 55-59	1	-0.3440	0.0526	42.94	< 0.0001
AGE13	Female, Age 60-64	1	-0.2240	0.0513	19.21	< 0.0001
AGE14	Female, Age 65-69	1	-0.2610	0.0520	25.30	< 0.0001

(CONTINUED)

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	-0.2300	0.0531	18.88	< 0.0001
AGE16	Female, Age 75-79	1	-0.1390	0.0548	6.51	0.0107
AGE17	Female, Age 80-84	1	-0.0170	0.0567	0.09	0.7587

c-statistic: Measures of association between the observed and predicted values were not calculated because the predicted probabilities are indistinguishable when they are classified into intervals of length 0.002.



**Table 14. Risk Adjustment Coefficients for PQI #90 Prevention Quality Overall Composite<sup>§</sup>**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-2.0360	0.0023	740186.3	< 0.0001
SEX	Female	1	-0.0760	0.0029	691.82	< 0.0001
AGE5	Male, Age 18-24	1	-4.1760	0.0062	445950.1	< 0.0001
AGE6	Male, Age 25-29	1	-3.9540	0.0066	349855.8	< 0.0001
AGE7	Male, Age 30-34	1	-3.7090	0.0061	361130.5	< 0.0001
AGE8	Male, Age 35-39	1	-3.4470	0.0055	385896.4	< 0.0001
AGE9	Male, Age 40-44	1	-3.0810	0.0047	423062.0	< 0.0001
AGE10	Male, Age 45-49	1	-2.7520	0.0041	447323.9	< 0.0001
AGE11	Male, Age 50-54	1	-2.4510	0.0037	422839.7	< 0.0001
AGE12	Male, Age 55-59	1	-2.2000	0.0036	363349.4	< 0.0001
AGE13	Male, Age 60-64	1	-1.9430	0.0035	297827.0	< 0.0001
AGE14	Male, Age 65-69	1	-1.5330	0.0035	192167.9	< 0.0001
AGE15	Male, Age 70-74	1	-1.1680	0.0034	112682.1	< 0.0001
AGE16	Male, Age 75-79	1	-0.8130	0.0034	55546.31	< 0.0001
AGE17	Male, Age 80-84	1	-0.4510	0.0034	17033.71	< 0.0001
AGE5	Female, Age 18-24	1	0.6521	0.0078	6879.14	< 0.0001
AGE6	Female, Age 25-29	1	0.4946	0.0086	3310.83	< 0.0001
AGE7	Female, Age 30-34	1	0.3677	0.0080	2066.28	< 0.0001
AGE8	Female, Age 35-39	1	0.3264	0.0072	2003.58	< 0.0001
AGE9	Female, Age 40-44	1	0.2465	0.0062	1539.85	< 0.0001
AGE10	Female, Age 45-49	1	0.2139	0.0054	1551.74	< 0.0001
AGE11	Female, Age 50-54	1	0.1674	0.0049	1134.66	< 0.0001
AGE12	Female, Age 55-59	1	0.1063	0.0048	485.41	< 0.0001
AGE13	Female, Age 60-64	1	0.1175	0.0046	631.07	< 0.0001
AGE14	Female, Age 65-69	1	0.0789	0.0045	296.44	< 0.0001

(CONTINUED)

<sup>§</sup> This PQI composite includes: PQIs #01, #03, #05, #07, #08, #10, #11, #12, #15 and #16. For more information see *Quality Indicator User Guide: Prevention Quality Indicators (PQI) Composite Measures*.

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	0.0660	0.0045	211.58	< 0.0001
AGE16	Female, Age 75-79	1	0.0460	0.0044	106.85	< 0.0001
AGE17	Female, Age 80-84	1	0.0147	0.0044	11.14	0.0008

c-statistic = 0.786

**Table 15. Risk Adjustment Coefficients for PQI #91\*\* - Prevention Quality Acute Composite**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-2.8470	0.0033	734218.8	< 0.0001
SEX	Female	1	0.0212	0.0040	27.43	< 0.0001
AGE5	Male, Age 18-24	1	-4.4240	0.0104	182328.4	< 0.0001
AGE6	Male, Age 25-29	1	-4.1960	0.0111	143502.4	< 0.0001
AGE7	Male, Age 30-34	1	-4.0050	0.0104	147307.9	< 0.0001
AGE8	Male, Age 35-39	1	-3.8010	0.0095	158222.6	< 0.0001
AGE9	Male, Age 40-44	1	-3.5430	0.0084	177710.9	< 0.0001
AGE10	Male, Age 45-49	1	-3.2420	0.0072	201576.7	< 0.0001
AGE11	Male, Age 50-54	1	-2.9360	0.0064	205351.9	< 0.0001
AGE12	Male, Age 55-59	1	-2.6620	0.0061	186272.4	< 0.0001
AGE13	Male, Age 60-64	1	-2.3450	0.0058	160698.9	< 0.0001
AGE14	Male, Age 65-69	1	-1.8740	0.0055	112438.1	< 0.0001
AGE15	Male, Age 70-74	1	-1.4280	0.0053	70136.33	< 0.0001
AGE16	Male, Age 75-79	1	-0.9810	0.0051	36038.17	< 0.0001
AGE17	Male, Age 80-84	1	-0.5450	0.0050	11749.08	< 0.0001
AGE5	Female, Age 18-24	1	0.9144	0.0123	5491.04	< 0.0001
AGE6	Female, Age 25-29	1	0.7361	0.0135	2985.49	< 0.0001
AGE7	Female, Age 30-34	1	0.6136	0.0129	2262.80	< 0.0001
AGE8	Female, Age 35-39	1	0.5503	0.0119	2138.03	< 0.0001
AGE9	Female, Age 40-44	1	0.4562	0.0106	1847.78	< 0.0001
AGE10	Female, Age 45-49	1	0.3563	0.0092	1493.51	< 0.0001
AGE11	Female, Age 50-54	1	0.2634	0.0083	995.51	< 0.0001
AGE12	Female, Age 55-59	1	0.1898	0.0079	564.09	< 0.0001
AGE13	Female, Age 60-64	1	0.1680	0.0075	492.82	< 0.0001
AGE14	Female, Age 65-69	1	0.0954	0.0072	173.44	< 0.0001

(CONTINUED)

\*\* This PQI composite includes: PQIs #10, #11 and #12. For more information see *Quality Indicator User Guide: Prevention Quality Indicators (PQI) Composite Measures*.

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	0.0660	0.0069	90.15	< 0.0001
AGE16	Female, Age 75-79	1	0.0315	0.0066	22.85	< 0.0001
AGE17	Female, Age 80-84	1	0.0225	0.0063	12.67	0.0004

c-statistic = 0.798

**Table 16. Risk Adjustment Coefficients for PQI #92<sup>††</sup> - Prevention Quality Chronic Composite**

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
INTERCEPT		1	-2.7410	0.0031	747523.5	< 0.0001
SEX	Female	1	-0.1640	0.0039	1719.42	< 0.0001
AGE5	Male, Age 18-24	1	-3.8990	0.0078	247809.8	< 0.0001
AGE6	Male, Age 25-29	1	-3.6800	0.0083	193305.6	< 0.0001
AGE7	Male, Age 30-34	1	-3.4080	0.0076	198267.2	< 0.0001
AGE8	Male, Age 35-39	1	-3.1190	0.0068	208526.5	< 0.0001
AGE9	Male, Age 40-44	1	-2.7090	0.0057	219448.4	< 0.0001
AGE10	Male, Age 45-49	1	-2.3700	0.0050	219292.1	< 0.0001
AGE11	Male, Age 50-54	1	-2.0730	0.0046	196236.0	< 0.0001
AGE12	Male, Age 55-59	1	-1.8330	0.0045	161711.9	< 0.0001
AGE13	Male, Age 60-64	1	-1.6030	0.0044	127710.8	< 0.0001
AGE14	Male, Age 65-69	1	-1.2260	0.0044	76235.05	< 0.0001
AGE15	Male, Age 70-74	1	-0.9120	0.0044	41673.93	< 0.0001
AGE16	Male, Age 75-79	1	-0.6240	0.0044	19349.02	< 0.0001
AGE17	Male, Age 80-84	1	-0.3320	0.0045	5351.63	< 0.0001
AGE5	Female, Age 18-24	1	0.4756	0.0103	2125.40	< 0.0001
AGE6	Female, Age 25-29	1	0.3364	0.0113	889.03	< 0.0001
AGE7	Female, Age 30-34	1	0.2271	0.0105	468.57	< 0.0001
AGE8	Female, Age 35-39	1	0.2233	0.0093	576.04	< 0.0001
AGE9	Female, Age 40-44	1	0.1811	0.0078	529.71	< 0.0001
AGE10	Female, Age 45-49	1	0.1923	0.0067	802.76	< 0.0001
AGE11	Female, Age 50-54	1	0.1689	0.0062	732.72	< 0.0001
AGE12	Female, Age 55-59	1	0.1114	0.0060	334.97	< 0.0001
AGE13	Female, Age 60-64	1	0.1327	0.0059	497.70	< 0.0001
AGE14	Female, Age 65-69	1	0.1068	0.0058	330.25	< 0.0001

(CONTINUED)

<sup>††</sup> This PQI composite includes: PQIs #01, #03, #05, #07, #08, #13, #14, #15 and #16. For more information see *Quality Indicator User Guide: Prevention Quality Indicators (PQI) Composite Measures*.

**AHRQ Quality Indicators™**  
**Prevention Quality Indicators (PQI) Parameter Estimates**

---

PARAMETER	LABEL	DF	ESTIMATE	STANDARD ERROR	WALD CHI-SQUARE	PR > CHI-SQUARE
AGE15	Female, Age 70-74	1	0.0959	0.0058	266.44	< 0.0001
AGE16	Female, Age 75-79	1	0.0777	0.0058	175.87	< 0.0001
AGE17	Female, Age 80-84	1	0.0210	0.0058	12.73	0.0004

c-statistic = 0.772

**Table A.1. Population Age Categories**

POPCAT	AGE RANGE
1	low - 4
2	5 - 9
3	10 - 14
4	15 - 17
5	18 - 24
6	25 - 29
7	30 - 34
8	35 - 39
9	40 - 44
10	45 - 49
11	50 - 54
12	55 - 59
13	60 - 64
14	65 - 69
15	70 - 74
16	75 - 79
17	80 - 84
18	85 - high



# **PREVENTION QUALITY INDICATORS (PQI) LOG OF ICD-9-CM AND DRG CODING UPDATES AND REVISIONS TO PQI DOCUMENTATION AND SOFTWARE Version 4.5**

**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. HHSA290201200001C**

**Prepared by:**

**Battelle**  
505 King Avenue  
Columbus, OH 43201

**May 2013**



## **Table of Contents**

1.0 Log of ICD-9-CM and DRG Coding Updates and Revisions to PQI Documentation and Software.....	1
---	---

## 1.0 Log of ICD-9-CM and DRG Coding Updates and Revisions to PQI Documentation and Software

The following table summarizes the revisions made to the Prevention Quality Indicator (PQI) software, software documentation and the technical specification documents since the original release of these documents in November 2001. It also reflects changes to indicator specifications based on updates to ICD-9-CM and MS-DRG codes through Fiscal Year 2013 (effective October 1, 2012) and incorporates coding updates that were implemented in both versions of the PQI software (both SAS and Windows).

The table lists the version and revision number, the date the revision was made, the component(s) affected by the change and a short summary of the changes that were made. The nature of the change is categorized into one of three types:

- 1) fiscal year (FY) coding change: occurs because of coding changes to the most recent fiscal year codes dictated by the Centers for Disease Control and Prevention (CDC) National Center for Health Statistics (NCHS) and Centers for Medicare & Medicaid Services (CMS),
- 2) specification/calculation change: may impact the measure result that is something other than the most recent fiscal year coding change, and
- 3) software/documentation change: alteration to the software code to calculate the measure as specified, or to the documentation to clarify the measure intent or functionality.

For convenience and ease of use, the changes are listed in reverse chronological order with the most recent changes appearing first in the table. Please note that changes prior to version 4.4 are not classified according to the currently defined types of changes. In addition, each type of change has varied shading to enhance readability.

All changes noted below have been incorporated into the software programming code, software documentation and the PQI technical specifications. With this software update, the PQI software now incorporates ICD-9-CM and DRG/MS-DRG codes valid from October 1, 1994 through September 30, 2013.

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.5	May 2013	All PQI	Specification/Calculation	Updated data are used for population estimates (i.e., through 2013). The population data are used to calculate the denominator for the area-level QI.
V4.5	May 2013	All PQI	Specification/Calculation	Updated reference population rates were calculated using 44 state files from the 2010 State Inpatient Databases (SID). New risk adjustment coefficients were calculated using the updated reference population.
V4.5	May 2013	Chronic Obstructive Pulmonary Disease (COPD) or Asthma in Older Adults Admission Rate (PQI 5)	Specification/Calculation	<p>Added numerator exclusion codes of any diagnosis of cystic fibrosis and anomalies of the respiratory system:</p> <p>27700 CYSTIC FIBROS W/O ILEUS  27701 CYSTIC FIBROS W ILEUS  27702 CYSTIC FIBROS W PUL MAN  27703 CYSTIC FIBROSIS W GI MAN  27709 CYSTIC FIBROSIS NEC  51661 NEUROENDOCRINE CELL HYPERPLASIA OF INFANCY  51662 PULMONARY INTERSTITIAL GLYCOGENESIS  51663 SURFACTANT MUTATIONS OF THE LUNG  51664 ALVEOLAR CAPILLARY DYSPLASIA WITH VEIN MISALIGNMENT  51669 OTHER INTERSTITIAL LUNG DISEASES OF THE CHILDHOOD  74721 ANOMALIES OF AORTIC ARCH  7483 LARYNGOTRACH ANOMALY NEC  7484 CONGENITAL CYSTIC LUNG  7485 AGENESIS OF LUNG  74860 LUNG ANOMALY NOS  74861 CONGEN BRONCHIECTASIS  74869 LUNG ANOMALY NEC  7488 RESPIRATORY ANOMALY NEC  7489 RESPIRATORY ANOMALY NOS  7503 CONG ESOPH FISTULA/ATRES  7593 SITUS INVERSUS  7707 PERINATAL CHR RESP DIS</p>

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.5	May 2013	All PQI	Software/Documentation	Respiratory complications diagnosis codes – Corrections were made to assure that three specific diagnosis codes were present in both the SAS and WinQI software. This change only affected the software.
V4.5	May 2013	All PQI	Software/Documentation	In WinQI there was an error in the smoothed rate calculation involving the noise variance and signal variance. This error was not previously observed because it only became significant in particular cases with relatively unusual variances. This issue was fixed in WinQI Version 4.5.
V4.5	May 2013	All PQI	Software/Documentation	The variable DISCWT in SAS QI v4.5 was set equal to 1 and the variable DISCWT was removed from the KEEP statement associated with the input file. This change ensures that the SAS programs do not account for complex sampling design when calculating QI estimates and standard errors. The SAS QI software, beginning with Version 4.1, does not support weighted QI estimates or standard errors for weighted estimates. The WinQI software has never supported weighted QI estimates or standard errors for weighted estimates.
V4.5	May 2013	All PQI	Software/Documentation	The installation packages have been improved for Version 4.5 of the SAS and WinQI software, including the Prediction Module and 3M™ APR DRG software. Both the SAS and WinQI software are available in Version 4.5 as either 32-bit or 64-bit applications. The 32-bit applications are targeted for Windows XP operating systems, and the 64-bit applications are targeted for Windows 7 operating systems.

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.5	May 2013	All PQI	Software/Documentation	<p>The WinQI software was corrected to address the following issues:</p> <ol style="list-style-type: none"> <li>1. On Step 2 of the Sampling Wizard dialog, the Sample Data File text box was not working correctly. Users were not able to save the file specified using the Browse explorer function. This issue has been fixed in WinQI Version 4.5.</li> <li>2. Denominators were not being adjusted (i.e., dividing by the number of discharge quarters) when the calculations were being stratified by quarter. This issue has been fixed in WinQI Version 4.5.</li> <li>3. On the WinQI Additional Options for Data Analysis screen of the Report Wizard, if the “<i>Ref. Pop. Rate</i>” is deselected, and then the expected rate and O/E ratio are reported incorrectly. These rates should be disabled on this screen if “<i>Ref. Pop. Rate</i>” is not selected. This issue has been included in the software documentation.</li> <li>4. The compiled C# program was named AHRQ.exe, and this was the same name used for the compiled Prediction Module C++ program. This potential conflict has been fixed in WinQI Version 4.5.</li> <li>5. Excel files with an .xlsx extension were not recognized. MS Access file types also needed to be updated. These issues were fixed in WinQI Version 4.5.</li> </ol>
V4.5	May 2013	Low Birth Weight Rate (PQI 9)	Software/Documentation	<ol style="list-style-type: none"> <li>1. A standalone SAS module was introduced that allows PQI #9 to be calculated without the need to run the entire PDI module. The PQI #9 Standalone Module for SAS is available for download from the AHRQ QI website.</li> <li>2. The definitions of Newborn and Outborn were revised in WinQI to better align them with SAS. The differences affected cases where discharge records have some combinations of missing values for one or more of the required data fields (e.g., Age, Age in Days).</li> </ol>

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.4	March 2012	All PQI	Specification/Calculation	Updated data are used for population estimates (i.e., through 2012). The population data are used to calculate the denominator for the area-level QI. The comparative data tables have been updated using Version 4.4 of the software. Because the risk adjustment models and reference population have not changed for Version 4.4, the Risk Adjustment Coefficients remain as they were in Version 4.3.
V4.4	March 2012	Hypertension Admission Rate (PQI 7)	Fiscal Year Coding	Add the following codes to existing numerator exclusions for cardiac procedures (PQI Appendix B)  Add code: 1755 TRANSLUM COR ATHERECTOMY 3505 ENDOVAS REPL AORTC VALVE 3506 TRANSPCL REP AORTC VALVE 3507 ENDOVAS REPL PULM VALVE 3508 TRANSAPCL REPL PULM VALVE 3509 ENDOVAS REPL UNS HRT VLV 3826 INSRT PRSR SNSR W/O LEAD
V4.4	March 2012	Heart Failure Admission Rate (PQI 8)		Add the following codes to existing numerator exclusions for cardiac procedures (PQI Appendix B)  Add code: 1755 TRANSLUM COR ATHERECTOMY 3505 ENDOVAS REPL AORTC VALVE 3506 TRANSPCL REP AORTC VALVE 3507 ENDOVAS REPL PULM VALVE 3508 TRANSAPCL REPL PULM VALVE 3509 ENDOVAS REPL UNS HRT VLV 3826 INSRT PRSR SNSR W/O LEAD

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.4	March 2012	Bacterial Pneumonia Admission Rate (PQI 11)	Fiscal Year Coding	Add exclusions for immunocompromised state diagnosis or procedures (PQI Appendix C)  Add code: 28411 ANTIN CHEMO INDCD PANCYT 28412 OTH DRG INDCD PANCYTOPNA 28419 OTHER PANCYTOPENIA 99688 COMP TP ORGAN-STEM CELL
V4.4	March 2012	Urinary Tract Infection Admission Rate (PQI 12)	Fiscal Year Coding	Add exclusions for immunocompromised state diagnosis or procedures (PQI Appendix C)  Add code: 28411 ANTIN CHEMO INDCD PANCYT 28412 OTH DRG INDCD PANCYTOPNA 28419 OTHER PANCYTOPENIA 99688 COMP TP ORGAN-STEM CELL
V4.4	March 2012	Angina without Procedure Admission Rate (PQI 13)	Fiscal Year Coding	Add exclusions for cardiac procedures (PQI Appendix B)  Add code: 1755 TRANSLUM COR ATHERECTOMY 3505 ENDOVAS REPL AORTC VALVE 3506 TRANSPCL REP AORTC VALVE 3507 ENDOVAS REPL PULM VALVE 3508 TRANSAPCL REPL PULM VALVE 3509 ENDOVAS REPL UNS HRT VLV 3826 INSRT PRSR SNSR W/O LEAD

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.4	March 2012	Asthma in Younger Adults Admission Rate (PQI 15)	Fiscal Year Coding	Add exclusions for cystic fibrosis and anomalies of respiratory system  Add code: 51661 NEUROEND CELL HYPRPL INF 51662 PULM INTERSTITL GLYCOGEN 51663 SURFACTANT MUTATION LUNG 51664 ALV CAP DYSP W VN MISALN 51669 OTH INTRST LUNG DIS CHLD
V4.4	March 2012	Heart Failure Admission Rate (PQI 8)	Software/Documentation	Rename indicator to Heart Failure Admission Rate  Rationale: Many patients with heart failure do not experience congestion of the lungs.
V4.4	March 2012	Software	Software/ Documentation	Revised the data step of creating permanent data set containing all records which are deleted from the analysis because key variable values having missing data
V4.4	March 2012	Software	Software/ Documentation	Both SAS and WinQI v4.3 were improperly truncating the (Observed rate)/ (Expected rate) ratio and associated upper confidence bound (95%) to be $\leq 1.0$ in cases where a stratification of the rates was being implemented. This issue was fixed in both SAS and WinQI so that this truncation only applies in cases where no stratification is being performed.
V4.4	March 2012	Software	Software/ Documentation	Sort routine (PROC SORT) was introduced to PQSASA3 programs before merging all the indicators together to sorting problems in SAS whenever user selects multiple stata (e.g. stratifies by age, gender, and age by gender)
V4.4	March 2012	Software	Software/ Documentation	PQSASA2.SAS program was revised to include denominator adjustment when the population count for certain combination of strata was zero.
V4.4	March 2012	Software	Software/ Documentation	WinQI v4.3 did not properly implement a user selection of year 2010 during report generation. This issue was fixed in v4.4 of WinQI.



VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.4	March 2012	Software	Software/ Documentation	WinQI v4.3 was not properly calculating quarterly rates when requested by the user. This issue was fixed in v4.4 of WinQI.
V4.4	March 2012	Software	Software/ Documentation	SAS v4.3 did not properly handle stratifications where the user requested a two-way stratification that overlapped with a one-way stratification (e.g., Age-by-Gender at the same time as Age by itself). This issue was fixed in v4.4 of SAS.
V4.4	March 2012	Software	Software/ Documentation	WinQI v4.3 and v4.4 do not check for a possible issue with user-defined composite weighting – users must set weights for all possible individual indicators, including zero weights for indicators that are not to be included in the composite. This requirement has been included in the software documentation.
V4.4	March 2012	Software	Software/ Documentation	SAS and WinQI v4.4 remain 32-bit applications developed on a Windows XP operating system. Some limited testing has been performed to ensure that these applications will run successfully under a 64-bit, Windows 7 environment. One additional installation requirement related to administrator rights has been included in the software documentation.
V4.4	March 2012	Software	Software/ Documentation	The software now incorporates state level estimates of diabetes prevalence by age from the CDC National Diabetes Surveillance System, which impacts PDI 15 and PQI 1, 3, 14, and 16.
V4.3	April 29, 2011	Hypertension Admission Rate (PQI 7) Numerator (Exclusion, cardiac procedure)	Coding	Add to numerator exclusion for cardiac procedure 3597 PERC MTRL VLV REPR W IMP 3737 EXC/DEST HRT LES, THRSPC

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.3	April 29, 2011	Congestive Heart Failure Admission Rate (PQI 8) Numerator (Exclusion, cardiac procedure)	Coding	Add to numerator exclusion for cardiac procedure 3597 PERC MTRL VLV REPR W IMP 3737 EXC/DEST HRT LES, THRSPC
V4.3	April 29, 2011	Angina Admission Rate (PQI 13) Numerator (Exclusion, cardiac procedure)	Coding	Add to numerator exclusion for cardiac procedure 3597 PERC MTRL VLV REPR W IMP 3737 EXC/DEST HRT LES, THRSPC
V4.3	April 29, 2011	Software (SAS and WinQI) and Documentation	Software/ Documents	PQI #5: Added numerator inclusion for principal diagnosis of asthma, modified numerator and denominator inclusion age to $\geq 40$ , and modified title to “Chronic Obstructive Pulmonary Disease or Asthma in Older Adults”
V4.3	April 29, 2011	Software (SAS and WinQI) and Documentation	Software/ Documents	PQI #10: Add numerator inclusion for secondary diagnosis of dehydration and principal diagnosis of hyperosmolality/hyponatremia, gastroenteritis, or acute renal failure. Added code for hyperosmolality/hyponatremia (276.0). Added numerator exclusion for chronic renal failure.
V4.3	April 29, 2011	Software (SAS and WinQI) and Documentation	Software/ Documents	PQI #15: Modified numerator and denominator inclusion to $\leq 40$ , modified title to “Asthma in Younger Adults”
V4.3	April 29, 2011	Software (SAS and WinQI) and Documentation	Software/ Documents	PQI #16: Added numerator exclusion for toe amputation (841.1)
V4.3	June 30, 2011	Software (SAS and WinQI) and Documentation	Software/ Documents	Surgical DRG: Added numerator inclusion codes 014 and 015 which were previously assigned to 009.

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.3	June 30, 2011	Guide	Software/ Documents	Revised and updated all sections of the guide document to reflect current state of indicators, software and body of evidence.
V4.2	September 30, 2010	Hypertension Admission Rate (PQI 7) Denominator (Exclusion)	Coding	Add procedure codes to denominator exclusion for Cardiac Procedures 17.51 Implantation of rechargeable cardiac contractility modulation (CCM), total system 17.52 Implantation or replacement of cardiac contractility modulation (CCM) rechargeable pulse generator only
V4.2	September 30, 2010	Bacterial Pneumonia Admission Rate (PQI 11) Denominator (Exclusion)	Coding	Add diagnosis codes to denominator exclusion for immunocompromised 279.41 Autoimmune lymphoproliferative syndrome ALPS 279.49 Autoimmune disease, not elsewhere classified
V4.1	December 2, 2009	SAS Software and Documentation	Software/ Documents	PQI #9 – Low Birth Weight – Added NOTE to documentation advising that this indicator is calculated by the PDI SAS module because it is based on pediatric discharges.
V4.0	June 30, 2009	Software and Documentation	Software/ Documents	PQI #7 – Hypertension – added numerator exclusion for diagnosis of Stage I-IV kidney disease only if accompanied by procedures for preparation for hemodialysis (dialysis access procedures)
V4.0	June 30, 2009	Software and Documentation	Software/ Documents	PQI #8 – CHF – dropped diagnosis codes from numerator inclusion for hypertension with heart disease and/or renal failure ONLY for discharges after 2002Q3 (effective Oct 1, 2002)
V4.0	June 30, 2009	Software and Documentation	Software/ Documents	PQI #11 – Bacterial pneumonia – added numerator exclusion for diagnosis code of immunocompromised state
V4.0	June 30, 2009	Software and Documentation	Software/ Documents	Cardiac procedure – added procedure codes to the numerator exclusion for cardiac procedures

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.0	June 30, 2009	SAS Software and Documentation	Software/ Documents	Implement UB-04 – The UB-04 (effective October 1, 2007) changes were implemented including new data elements for point-of-origin and present on admission
V4.0	June 30, 2009	SAS Software and Documentation	Software/ Documents	Update Benchmarking Data to 2007 – used data from the 2007 SID for computation of benchmarks
V4.0	February 20, 2009	Bacterial Pneumonia Admission Rate (PQI 11) Numerator (Inclusion)	Coding	Add diagnosis code to numerator inclusion for bacterial pneumonia (\$ACSBACD) Modify code: 482.41 Methicillin susceptible pneumonia due to Staphylococcus aureus Add code: 482.42 Methicillin resistant pneumonia due to Staphylococcus aureus
V4.0	February 20, 2009	Cardiac procedures	Coding	Add procedure codes to numerator exclusion for cardiac procedures (\$ACSCARP) Add codes: 37.36 Excision or destruction of left atrial appendage (LAA) 37.55 Removal of internal biventricular heart replacement system 37.60 Implantation or insertion of biventricular external heart assist system
V4.0	February 20, 2009	Immunocompromis ed	Coding	Add diagnosis codes to numerator exclusion for immunocompromised (\$IMMUNID) 199.2 Malignant neoplasm associated with transplanted organ 238.77 Neoplasm of uncertain behavior, post-transplant lymphoproliferative disorder (PTLD) 238.79 Neoplasm of uncertain behavior, other lymphatic and hematopoietic tissues 279.50 Graft-versus-host disease unspecified 279.51 Acute graft-versus-host disease 279.52 Chronic graft-versus-host disease 279.53 Acute on chronic graft-versus-host disease V45.11 Renal dialysis status

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.0	February 20, 2009	Hypertension Admission Rate (PQI 7) Numerator (Exclusion)	Indicator Specification	<p>Add numerator exclusion for diagnosis of Stage I-IV kidney disease (\$ACSHY2D) only if accompanied by procedures for preparation for hemodialysis (dialysis access procedures) (\$ACSHYPP).</p> <p>Add codes:</p> <p>403.00 Hypertensive chronic kidney disease, malignant, with chronic kidney disease stage I through stage IV, or unspecified</p> <p>403.10 Hypertensive chronic kidney disease, benign, with chronic kidney disease stage I through stage IV, or unspecified</p> <p>403.90 Hypertensive chronic kidney disease, unspecified, with chronic kidney disease stage I through stage IV, or unspecified</p> <p>404.00 Hypertensive heart and chronic kidney disease, malignant, without heart failure and with chronic kidney disease stage I through stage IV, or unspecified</p> <p>404.10 Hypertensive heart and chronic kidney disease, benign, without heart failure and with chronic kidney disease stage I through stage IV, or unspecified</p> <p>404.90 Hypertensive heart and chronic kidney disease, unspecified, without heart failure and with chronic kidney disease stage I through stage IV, or unspecified</p> <p>ONLY if codes:</p> <p>38.95 Venous catheterization for renal dialysis</p> <p>39.27 Arteriovenostomy for renal dialysis</p> <p>39.29 Other (peripheral) vascular shunt or bypass</p> <p>39.42 Revision of arteriovenous shunt for renal dialysis</p> <p>39.43 Removal of arteriovenous shunt for renal dialysis</p> <p>39.93 Insertion of vessel-to-vessel cannula</p> <p>39.94 Replacement of vessel-to-vessel cannula</p>

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.0	February 20, 2009	Congestive Heart Failure Admission Rate (PQI 8) Numerator (Inclusion)	Indicator Specification	<p>Drop diagnosis codes from numerator inclusion for hypertension with heart disease and/or renal failure (\$ACSCH2D) ONLY for discharges after 2002Q3 (effective Oct 1, 2002)</p> <p>Delete codes:</p> <p>402.01 Hypertensive heart disease, malignant, with heart failure</p> <p>402.11 Hypertensive heart disease, benign, with heart failure</p> <p>402.91 Hypertensive heart disease, unspecified, with heart failure</p> <p>404.01 Hypertensive heart and chronic kidney disease, malignant, with heart failure and with chronic kidney disease stage I through stage IV, or unspecified</p> <p>404.03 Hypertensive heart and chronic kidney disease, malignant, with heart failure and with chronic kidney disease stage V or end stage renal disease</p> <p>404.11 Hypertensive heart and chronic kidney disease, benign, with heart failure and with chronic kidney disease stage I through stage IV, or unspecified</p> <p>404.13 Hypertensive heart and chronic kidney disease, benign, with heart failure and chronic kidney disease stage V or end stage renal disease</p> <p>404.91 Hypertensive heart and chronic kidney disease, unspecified, with heart failure and with chronic kidney disease stage I through stage IV, or unspecified</p> <p>404.93 Hypertensive heart and chronic kidney disease, unspecified, with heart failure and chronic kidney disease stage V or end stage renal disease</p>
V4.0	February 20, 2009	Bacterial Pneumonia Admission Rate (PQI 11) Numerator (Exclusion)	Indicator Specification	Add numerator exclusion for diagnosis code of immunocompromised state (\$IMMUNIP)

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V4.0	February 20, 2009	Cardiac Procedure	Indicator Specification	Add procedure codes to the numerator exclusion for cardiac procedures (\$ACSCARP) Add codes: 37.61 Implant of pulsation balloon 37.62 Insertion of non-implantable heart assist system 37.63 Repair of heart assist system 37.64 Removal of heart assist system 37.65 Implant of external heart assist system 37.66 Insertion of implantable heart assist system
V3.2	March 10, 2008		Coding	There were no changes to ICD-9-CM or DRG codes
V3.2	March 10, 2008	None	Software/ Documents	No change to software or documents
V3.1a	March 16, 2007	SAS Software (PQSASA2)	Software/ Documents	Amended the aggregation algorithm to correctly sum the numerator and denominator counts across stratifiers.
V3.1	March 12, 2007	Software (SAS and Windows), Software Documentation, Guide, and Technical Specifications	Software/ Documents	Implemented changes associated with ICD-9-CM coding updates for Fiscal Year (FY) 2007 (effective 10-1-2006). See separate documentation on ICD-9 coding updates for specific details. The years for which the ICD-9-CM and DRG codes defining PQIs are valid was amended to be through FY 2007 instead of FY 2006, that is, the codes in the software are effective through September 30, 2007.
V3.1	March 12, 2007	Covariates. Software (SAS and Windows)	Software/ Documents	Based on recommendations of the Risk Adjustment and Hierarchical Modeling (RAHM) Workgroup, computed covariates using a logistic regression model with an area random-effect instead of the existing simple logistic model. Because the AHRQ QI use a “large sample”, the impact on the covariates of using the hierarchical model (and hence the impact on the risk-adjusted rates) is minor.

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V3.1	March 12, 2007	Software (SAS and Windows), Software Documentation and Covariates document	Software/ Documents	Updated the coefficients used in the calculation of expected and risk-adjusted rates to the 2002-2004 reference population.
V3.1	March 12, 2007	Technical Specifications	Software/ Documents	Moved list of ICD-9-CM codes for cardiac procedure into an Appendix, with links to and from the PQIs that use the codes as a numerator exclusion.
V3.1	March 12, 2007	Guide	Software/ Documents	Moved average volume, provider rates, and population rates into separate document, <i>Prevention Quality Indicators Comparative Data</i>
V3.1	March 12, 2007	Software (SAS and Windows)	Software/ Documents	Age-, race-, gender- and county-specific population estimates used for AHRQ QI area rates were updated to use revised post-censal estimates for years 2001 through 2005 and projections for the years 2006 and 2007. Modified the A3 syntax to compute risk-adjusted rates and observed-to-expected ratios for the pre-defined set of stratification variables (e.g., age, gender, payer, race) Added option to select whether or not to apply county-level adjustment for Socioeconomic Status (SES) and/or disease prevalence in addition to age and gender.
V3.0b	May 1, 2006	Technical Specifications	Software/ Documents	Revised denominator description for PQI #9. Deleted codes 59000 and 59001 from numerator of PQI #10. Corrected code numbers in denominator of PQI #13.
V3.0b	May 1, 2006	All documents	Software/ Documents	Edited PDF files to make URLs in header or footnotes clickable links.
V3.0a	February 20, 2006	Hypertension Admission Rate (PQI 7) (Exclusion)	Coding	Added new (FY2006) codes 00.66 "Percutaneous Transluminal Coronary Angioplasty" and 37.41 "Implantation of prosthetic cardiac support device around the heart" to the cardiac procedure exclusion.



VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V3.0a	February 20, 2006	Congestive Heart Failure Admission Rate (PQI 8) (Exclusion)	Coding	Added new (FY2006) codes 00.66 "Percutaneous Transluminal Coronary Angioplasty" and 37.41 "Implantation of prosthetic cardiac support device around the heart" to the cardiac procedure exclusion.
V3.0a	February 20, 2006	Dehydration (PQI 10) Numerator	Coding	Added new (FY2006) codes 276.50 "Volume depletion, unspecified", 276.51 "Dehydration", and 276.52 "Hypovolemia" to the inclusion criteria.
V3.0a	February 20, 2006	Urinary Tract Infection (PQI 12) Numerator (Exclusion)	Coding	Added exclusion for any diagnosis code of kidney/urinary tract disorder and for any diagnosis code of immunocompromised state.
V3.0a	February 20, 2006	Angina without Procedure Admission Rate (PQI 13) (Exclusion)	Coding	Added new (FY2006) codes 00.66 "Percutaneous Transluminal Coronary Angioplasty" and 37.41 "Implantation of prosthetic cardiac support device around the heart" to the cardiac procedure exclusion.
V3.0a	February 20, 2006	Asthma (PQI 15) Numerator (Exclusion)	Coding	Added exclusion for any diagnosis code of cystic fibrosis and anomalies of the respiratory system.
V3.0a	February 20, 2006	Guide, SAS and SPSS Software Documentation	Software/ Documents	Removed Appendices that were copies of Change Log and Indicator Changes documents. Added Appendix of Links to all PQI documents and additional resources.
V3.0a	February 20, 2006	Guide	Software/ Documents	Added explanation of changes to area definitions and new stratification options. Changed "MSA" to "Metro Area" throughout the document. Added section "Using Different Types of QI rates."

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V3.0a	February 20, 2006	Software, Guide, and Technical Specifications	Software/ Documents	Revised denominator of PQI #9 (Low Birth Weight) to define newborn as neonate with age at admission of 0 to 28 days, with ICD-9-CM diagnosis code for in-hospital live birth. Revised numerator of PQI #12 (Urinary Tract Infection) to Add exclusion for any diagnosis code of kidney/urinary tract disorder and for any diagnosis code of immunocompromised state. Revised numerator of PQI #15 (Asthma) to Add exclusion for any diagnosis code of cystic fibrosis and anomalies of the respiratory system.
V3.0a	February 20, 2006	Software (SAS and SPSS) Software Documentation	Software/ Documents	Changed name of data element HOSPSTCO to PSTCO. Added parameter POPYEAR to specify year for Census data. Changed name of MSALEVL parameter to MALEVL to reflect the change in OMB definitions for areas, and added options to allow users to specify stratification by county level with U.S. Census FIPS or modified FIPS, or Metro Area with OMB 1999 or OMB 2003 definition.
V3.0a	February 20, 2006	Software (SAS and SPSS)	Software/ Documents	Changed the computation of the risk-adjusted rate to use a proportional formula for indirect standardization.
V3.0a	February 20, 2006	Software (SAS)	Software/ Documents	Added a computation of confidence limits.
V3.0a	February 20, 2006	Software (SAS and SPSS), Software Documentation and Covariates document	Software/ Documents	Updated the coefficients used in the calculation of expected and risk-adjusted rates to the 2003 reference population.
V3.0a	February 20, 2006	Indicator Changes	Software/ Documents	Revised to limit entries to indicator changes made because of changes to ICD-9-CM code updates for FY2006 and moved entries for specification changes into PQI Change Log.

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V3.0	November 30, 2005	Guide	Software/ Documents	Moved Appendix A into new document <i>Prevention Quality Indicators Technical Specifications</i> . Removed Appendix B.
V3.0	November 30, 2005	Software (SAS and SPSS), Software Documentation, Guide, Technical Specifications, and Analysis & Interpretation	Software/ Documents	Implemented changes associated with ICD-9-CM coding updates for Fiscal Year (FY) 2006 (effective 10-1-2005). See separate documentation on ICD-9 coding updates for specific details. The years for which the ICD-9-CM and DRG codes defining PQIs are valid was amended to be through FY 2006 instead of FY 2005, that is, the codes in the software are effective through September 30, 2006. Dropped PQI #4 and PQI #6, which are being moved into the new Pediatric Quality Indicators module. Revised PQI #2, PQI #10, PQI #11, and PQI #12 to exclude pediatric populations. Added exclusion for cystic fibrosis and anomalies of the respiratory system to PQI #15 (Asthma). Added exclusion for kidney/urinary tract disorder and immunocompromised state to PQI #12 (Urinary Tract Infection).
V3.0	November 30, 2005	Software Documentation (SAS and SPSS)	Software/ Documents	Removed section "Interpreting the Results." Table 3 was amended to include the 2004-06 census data and condition-specific module file (i.e., QICTYCyy.TXT).
V3.0	November 30, 2005	Software (SAS and SPSS)	Software/ Documents	Added the 2004-06 census data and condition-specific module file (e.g., QICTYCyy.TXT)
V2.1 R4	November 24, 2004		Coding	There were no ICD-9-CM or DRG coding changes that affected indicator definitions.
V2.1 R4	November 24, 2004	Software (SAS and SPSS), Software Documentation, and Guide	Software/ Documents	The years for which the ICD-9-CM and DRG codes defining PQIs are valid was amended to be through FY 2005 instead of FY 2004, that is, the codes in the software are effective through September 30, 2005. Added new module that calculates condition-specific rates for the diabetes PQIs across stratifiers.

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V2.1 R4	November 24, 2004	Software Documentation (SAS and SPSS)	Software/ Documents	Table 3 was amended to include the 2003 census data (i.e., QICTY03.TXT and QICTYA03.TXT) and condition-specific module files (PQSASC2 and QICTYC03.TXT).
V2.1 R4	November 24, 2004	Software (SAS and SPSS)	Software/ Documents	Added the 2003 census data (i.e., QICTY03.TXT and QICTYA03.TXT) and condition-specific module files (PQSASC2 and QICTYC03.TXT)
V2.1 R4	November 24, 2004	Guide	Software/ Documents	Rearranged the sequence of PQIs to place in numerical order.
V2.1 R4	November 24, 2004	Software (SAS)	Software/ Documents	Inserted “PQ” in format names for age aggregations in SAS programs to distinguish these formats from similarly named formats used by other indicator software.
V2.1 R3	January 9, 2004	Bacterial Pneumonia Admission Rate (PQI 11) Numerator (Exclusion, sickle cell anemia and HB-S disease)	Coding	New codes (FY 2004) 282.41, 282.42, 282.64, 282.68 were added to the numerator exclusion definition of HB-S and sickle cell anemia. This change may result in a comparability issue with previous years since 282.4 was not previously included in the sickle cell definition.
V2.1 R3	January 9, 2004	Adult Asthma Admission Rate (PQI 15) Numerator	Coding	New codes (FY 2004), 493.81 “Exercised Induced Bronchospasm” and 493.82 “Cough Variant Asthma” were added to the numerator definition of asthma
V2.1 R3	January 9, 2004	Pediatric Asthma Admission Rate (PQI 4) Numerator	Coding	New codes (FY 2004), 493.81 “Exercised Induced Bronchospasm” and 493.82 “Cough Variant Asthma” were added to the numerator definition of asthma
V2.1 R3	January 9, 2004	Congestive Heart Failure Admission Rate (PQI 8) Numerator	Coding	The new codes (FY 2003), 428.20-3, “Systolic heart failure,” 428.30-3, “Diastolic heart failure,” and 428.40-3, “Combined systolic and diastolic heart failure” were added to the including definition of congestive heart failure.

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V2.1 R3	January 9, 2004	Congestive Heart Failure Admission Rate (PQI 8) Numerator (Exclusion, cardiac procedures)	Coding	<p>The new code (FY 2003), 36.07, “Insertion of drug-eluting coronary artery stent(s) Endograft(s), Endovascular graft(s), Stent graft(s)” was added to the exclusion definition of cardiac procedures.</p> <p>The new codes (FY 2003), 00.50-00.54, “implantation or replacement of transvenous lead” were added to the exclusion definition of cardiac procedures.</p> <p>All new codes (FY 2004) in the new category heart replacement procedures (37.5), including 37.51, “heart transplantation,” 37.52 “implantation of total replacement heart system,” 37.53 “replacement or repair of thoracic unit of total replacement heart system,” and 37.54 “replacement or repair of other implantable component of total replacement heart system” were added to the numerator exclusion definition of cardiac procedure. Note that 37.5, previously used for heart transplantation procedure is invalid as of October 2003. This code was retained in the software for backward comparability.</p>
V2.1 R3	January 9, 2004	Hypertension Admission Rate (PQI 7) Numerator (Exclusion, cardiac procedures)	Coding	<p>The new code (FY 2003), 36.07, “Insertion of drug-eluting coronary artery stent(s) Endograft(s), Endovascular graft(s), Stent graft(s)” was added to the exclusion definition of cardiac procedures.</p> <p>The new codes (FY 2003), 00.50-00.54, “implantation or replacement of transvenous lead” were added to the exclusion definition of cardiac procedures.</p> <p>All new codes (FY 2004) in the new category heart replacement procedures (37.5), including 37.51, “heart transplantation,” 37.52 “implantation of total replacement heart system,” 37.53 “replacement or repair of thoracic unit of total replacement heart system,” and 37.54 “replacement or repair of other implantable component of total replacement heart system” were added to the numerator exclusion definition of cardiac procedure. Note that 37.5, previously used for heart transplantation procedure is invalid as of October 2003. This code was retained in the software for backward comparability.</p>

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V2.1 R3	January 9, 2004	Angina Admission Rate (PQI 13) Numerator (Exclusion, cardiac procedures)	Coding	<p>The new code (FY 2003), 36.07, “Insertion of drug-eluting coronary artery stent(s) Endograft(s), Endovascular graft(s), Stent graft(s)” was added to the exclusion definition of cardiac procedures.</p> <p>The new codes (FY 2003), 00.50-00.54, “implantation or replacement of transvenous lead” were added to the exclusion definition of cardiac procedures.</p> <p>All new codes (FY 2004) in the new category heart replacement procedures (37.5), including 37.51, “heart transplantation,” 37.52 “implantation of total replacement heart system,” 37.53 “replacement or repair of thoracic unit of total replacement heart system,” and 37.54 “replacement or repair of other implantable component of total replacement heart system” were added to the numerator exclusion definition of cardiac procedure. Note that 37.5, previously used for heart transplantation procedure is invalid as of October 2003. This code was retained in the software for backward comparability.</p>
V2.1 R3	January 9, 2004	Software (SAS and SPSS) and Guide	Software/ Documents	Implemented changes associated with ICD-9-CM coding updates from Fiscal Year (FY) 2003 (effective 10-1-2002) and FY 2004 (effective 10-1-2003). See separate documentation on ICD-9 coding updates for specific details.

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V2.1 R3	January 9, 2004	Software (SAS and SPSS)	Software/ Documents	Angina Admission Rate. The numerator exclusion for patients undergoing any surgical procedure was removed and replaced with a more restrictive exclusion of cardiac procedures, identical to the exclusion list for cardiac procedures included in the CHF Admission Rate and Hypertension Admission Rate Indicators (see below). The rate for the Angina Admission Rate indicator is expected to decrease significantly with this change. CHF Admission Rate, Hypertension Admission rate and Angina Admission Rate. The numerator exclusion of major cardiac surgery was redefined to include only surgeries that would typically be done on an elective or semi-elective basis and therefore represent the indication for admission. This would include valve repair (35.xx), angioplasty and stent placement (36.0x), coronary bypass and other revascularization surgery (36.1x-36.9x), and heart transplantation (37.5). In addition, the list was expanded to include procedures associated with angina, in conjunction with the use of this inclusion in the Angina Admission Rate Indicator. The resulting exclusion is now identical for the three indicators.
V2.1 R3	January 9, 2004	Software (SAS and SPSS)	Software/ Documents	All parameter text files were renamed to refer specifically to the PQI module (e.g., use of PQ in file name). These changes are also reflected in the software documentation. All parameter files were rerun using the updated software and Year 2000 HCUP SID data. Population files for 2000, 2001 and 2002 were re-estimated using the latest available census files
V2.1 R3	January 9, 2004	Software – SPSS	Software/ Documents	The treatment of missing data by SPSS was changed to mirror the treatment of missing data by SAS, specifically the software requires confirmation for the assignment of a poor outcome or negative event. For instance, in order to be assigned as a death, each case must actually be coded as a death. Missing data is considered neutral. Missing data for some elements results in the exclusion of that case from the denominator. For a few other elements, the case is retained. Table 5 of the Software Documentation lists the impact of missing data for each data element.

VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V2.1 R2	January 10, 2003	Software Documentation (SAS and SPSS)	Software/ Documents	Updated documentation to reference the changes made to the software programs such as the change in the default number of ICD-9 diagnosis and procedure codes, the option to stratify area by MSA or county, and instructions for using the patient FIPS code. Modified the data file input specifications to standardize across software programs (SAS and SPSS) so the user would be able to run the same input data file with either statistical package.
V2.1 R2	January 10, 2003	Software (SAS and SPSS)	Software/ Documents	The county-to-MSA mapping for Waller County in Texas was corrected by assigning the value of 3362 for the Houston-Galveston MSA. The default number of ICD-9-CM diagnoses was changed from 5 to 30. The default number of ICD-9-CM procedures was changed from 4 to 30. The ICD-9 coding was updated to reflect changes through FY 2002 (September 30, 2002). Added the option for the user to select rates calculated by MSA or by county for urban areas (rates for rural areas will always be by county). Additional ASCII text files with Census residential population numbers for 2000 and 2001 were included in the module. Risk-adjustment inputs that were based on nineteen SID state data files from the year 1997 were replaced with numbers that were based on twenty-nine SID state data files from the year 2000. The formulation of smoothed rates was corrected so that missing values would be generated when appropriate, rather than zeros. Hardcopy printouts were modified to be easier to understand (intermediate means were removed, the final means were restricted to just area-level records, prints of the final results were reformatted and labeled).
V2.1 R2	October 9, 2002	Guide	Software/ Documents	The definition for the Perforated appendix admission rate was clarified in appendix A, by moving the ICD-9-CM codes for the population at risk to a separate section that defined the denominator for the rate. The definition of the Low Birthweight indicator was corrected in Appendix A, by removing references to DRG's 370-375.



VERSION/ REVISION NUMBER	DATE	COMPONENT	NATURE OF CHANGE	CHANGES
V2.1 R1	April 17, 2002	Guide	Software/ Documents	<p>The age inclusions for the populations at risk were corrected for the following indicators: bacteria pneumonia, dehydration, urinary tract infection, angina without procedure, CHF, hypertension, adult asthma, COPD, uncontrolled diabetes, diabetes short-term complications, diabetes long-term complications, and lower-extremity amputation among patients with diabetes. In all cases, the descriptions of the indicators in the Guide suggested that the indicator be applied to a specific age group, but suggested that it could be applied to other age groups as well. The software applies the indicator to all relevant age groups; therefore, the Guide was amended to reflect this.</p> <p>For the definition of Lower-Extremity Amputation among Patients with Diabetes, under Outcomes of Interest, "Discharges with ICD-9-CM <i>principal diagnosis</i> codes" was changed to "Discharges with ICD-9-CM <i>procedure</i> codes".</p>
V2.1 R1	April 17, 2002	Software documentation	Software/ Documents	<p>The years for which the ICD-9-CM codes defining PQIs are valid was amended to be through FY 2001 instead of FY 2000, that is, the codes in the software are effective through September 30, 2001.</p>



## 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. HHSA290201200001C**

**Prepared by:**

**Battelle**

505 King Avenue  
Columbus, OH 43201

**Revised November 2013**

## Disclaimer

This document is based on research conducted 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:**

Geppert, J., Rhoda, D., Morara, M. Quality Indicator Empirical Methods (Revised by Battelle, under Contract No. HHSA290201200001C). Rockville, MD: Agency for Healthcare Research and Quality. September 2013.

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

## Contents

Overview .....	2
Chapter 1. The User's Dataset .....	5
Chapter 2. Calculating Volume and Count Indicators .....	7
Chapter 3. Calculating Area-Level Indicators – Observed Rates .....	9
Chapter 4. Risk Adjustment for Area-Level Indicators .....	12
Chapter 5. Calculating Area-Level Indicators – Expected, Risk-Adjusted, & Smoothed Rates ..	13
Chapter 6. Overview of Provider-Level QI & Present on Admission (POA) .....	15
Chapter 7. Calculating Provider-Level Observed Rates – Ignoring POA .....	19
Chapter 8. Calculating Provider-Level Observed Rates – With Complete POA Data .....	21
Chapter 9. Calculating Provider-Level Observed Rates – With Missing POA Data.....	23
Chapter 10. Risk Adjustment for Provider-Level Indicators .....	26
Chapter 11. Estimating Composite Measures .....	34
Chapter 12. Software Maintenance – Updating the Reference Population .....	39
Chapter 13. Software Maintenance – Other Updates.....	44
References .....	45
Appendix A. Table of AHRQ QI Risk-adjustment / POA.....	47
Appendix B. Table of AHRQ QI Provider-Level Risk-adjustment Covariates.....	52
Appendix C. Provider-Level Risk Adjustment - Detailed Methods .....	54
Appendix D. Helpful Background Information .....	65

## List of Tables

Table 1.1 Required Data Elements .....	5
Table 1.2 Data Elements and Data Values To Be Constructed by the User .....	6
Table 1.3 Analysis Data Inclusion Rule .....	6
Table 2.1 AHRQ QI Volume Indicators .....	7
Table 2.2 AHRQ QI Count Indicators .....	7
Table 3.1 AHRQ QI Area-Level Indicators.....	9
Table 6.1 Methods Used by QI Software to Distinguish Complications from Comorbidities ....	16
Table 6.2 Values for the Present-on-Admission Data Element .....	18
Table 10.1 Parameter Estimates CSV Files for Provider-Level Risk Adjustment .....	27

## Overview

This document describes the empirical methods used to calculate the AHRQ Quality Indicators™ (AHRQ QI). The QI 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](http://www.qualityindicators.ahrq.gov).

The current AHRQ QI modules represent various aspects of quality:

- Prevention Quality Indicators (PQI) identify hospital admissions in geographic areas that evidence suggests might have been avoided through access to high-quality outpatient care. (first released November 2000, last updated May 2013)
- Inpatient Quality Indicators (IQI) 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 May 2013)
- Patient Safety Indicators (PSI) 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 May 2013)
- Pediatric Quality Indicators (PDI) 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 identify potentially avoidable hospitalizations. (first released April 2006, last updated May 2013)

The input data for QI calculation consist of 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 healthcare provider's facility?) or **area-level** (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?). Some indicators report the number of times a hospital performed a medical **procedure of interest**. These **volume indicators** do not have denominators. Most of the AHRQ QI are ratios where 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 several rates:

1. **Observed rate** – Conceptually, provider-level rates are the number of discharge records where 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 persons who live in that area who are at risk for the condition.
2. **Expected rate** – A comparative rate that incorporates information about a **reference population** that is not part of the user's input dataset – what rate would be observed 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.

3. **Risk-adjusted rate** – A comparative rate that also incorporates information about a **reference population** that is not part of the input dataset – what rate 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 which QIs are risk-adjusted.
4. **Smoothed rate** – A weighted average of the risk-adjusted rate from the user's input dataset and the rate observed in the **reference population**; the smoothed rate is calculated with a shrinkage estimator 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 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 mean, and does this more so for outliers (such as rural hospitals).

In data collected beginning October 1, 2007, each diagnosis code may be accompanied by a data element that indicates whether the diagnosed condition was **Present-on-Admission (POA)**, and is therefore a pre-existing **comorbidity**, or whether the condition developed during the hospitalization of interest and is therefore a **complication**. Some datasets include POA data, while others do not. Some datasets have POA data for many, but not all of the discharge records. POA is handled in different ways in the QI software depending on a) whether POA data are present in the discharge record and b) whether the user specifies that the software should use the POA data elements when calculating QI rates, or ignore the POA data elements.

This document begins with a brief description of the dataset that a user must assemble to run the QI software and then it describes the methods associated with various types of indicators. Simpler indicators are described first. Volume indicators are the simplest of the QI. 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 the calculation of provider-level indicators, where the denominator is tailored to the indicator and the QI may be affected by the POA data element, and how the software accounts for missing POA data. Composite indicators are described next and then the document finishes 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. Helpful examples include:

### 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](http://www.qualityindicators.ahrq.gov/Modules/PQI_TechSpec.aspx)  
IQI: See [http://www.qualityindicators.ahrq.gov/Modules/IQI\\_TechSpec.aspx](http://www.qualityindicators.ahrq.gov/Modules/IQI_TechSpec.aspx)  
PSI: See [http://www.qualityindicators.ahrq.gov/Modules/PSI\\_TechSpec.aspx](http://www.qualityindicators.ahrq.gov/Modules/PSI_TechSpec.aspx)  
PDI: See [http://www.qualityindicators.ahrq.gov/Modules/PDI\\_TechSpec.aspx](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](http://www.qualityindicators.ahrq.gov/modules/pqi_resources.aspx)  
IQI: See [http://www.qualityindicators.ahrq.gov/modules/iqi\\_resources.aspx](http://www.qualityindicators.ahrq.gov/modules/iqi_resources.aspx)  
PSI: See [http://www.qualityindicators.ahrq.gov/modules/psi\\_resources.aspx](http://www.qualityindicators.ahrq.gov/modules/psi_resources.aspx)  
PDI: See [http://www.qualityindicators.ahrq.gov/modules/pdi\\_resources.aspx](http://www.qualityindicators.ahrq.gov/modules/pdi_resources.aspx)

#### QI Population Documentation File

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

#### QI Prediction Module Testing Report

See <http://www.qualityindicators.ahrq.gov/Modules/Default.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-DXPOA30	Diagnosis present on admission indicator		X	X	X
E_POA1-E_POA10	E code present on admission indicator		X	X	X
ECODE1-ECODE10	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
POINTOFORIGINUB04	Point of origin for admission or visit, UB-04 standard coding	X	X	X	X
PR1-PR30	Procedure	X	X	X	X
PRDAY1-PRDAY30	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 ZIP Code	X	X	X	X
RACE	Race (uniform)	X	X	X	X
SEX	Sex	X	X	X	
YEAR	Calendar year	X	X	X	X

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 in 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

Discharge records in the dataset are analyzed as either adult or pediatric data based on 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 greater than or equal to 18 or MDC equal to 14
Pediatric	AGE less than 18 and MDC not equal to 14

Adult analysis data are used to calculate Prevention Quality Indicators (PQI), Inpatient Quality Indicators (IQI), and Patient Safety Indicators (PSI). Pediatric records are used to calculate Pediatric Quality Indicators (PDI), Neonatal Quality Indicators (NQI) 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

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 05 – 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 phrases **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 phrase “numerator” to define the procedure of interest. Discharge records are flagged for inclusion or exclusion from the numerator of each volume QI based on the data elements, data values, and logic described in the technical specifications for each indicator.

For each discharge record, a binary flag variable is calculated by the software for each volume or count QI. In this document, we denote the discharge level indicator data element with the letter T. Each discharge record has a T variable for each QI, so in the software the data elements have longer names to clarify which QI they describe. (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. Discharges are also excluded from the numerator of a volume QI if the procedure of interest has more than one component, and the discharge is not in the population at risk for one component but remains in the population at risk for another component. These discharges are assigned a value of “0” for T.

## **The Observed Value**

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

## Chapter 3. Calculating Area-Level Indicators – Observed Rates

**Area-level indicators** identify hospital admissions that evidence suggests might have been avoided through access to high-quality outpatient or preventive care. 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.

Table 3.1 lists the area level indicators.

**Table 3.1 AHRQ QI 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 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 (CMSAs), and primary metropolitan statistical areas (PMSAs). In addition, “area” could refer to

either 1) FIPS county, 2) modified FIPS county, 3) 1999 OMB Metropolitan Statistical Area, or 4) 2003 OMB Metropolitan Statistical Area. 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 where 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 3 Diabetes Long-Term Complications Admission Rate, PQI 14 Uncontrolled Diabetes Admission Rate, and PQI 16 Lower-Extremity Amputation among Patients with Diabetes Rate. [Chapter 13](#) 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

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, and therefore not an indication of substandard care.
2. The patient was transferred from another health care facility.
3. Some exclusion criteria are included for the purpose of enhancing face validity with clinicians.
4. Some exclusion criteria are an inherent part of the QI definition.

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

## The Observed Rate

The observed rate of an area-level indicator is simply the sum of T 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.

## 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** which 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 gender, and optionally for poverty. That is to say that the models include age (in 5 year groups), gender, and if it is statistically significant, the model includes the interaction between age and gender.

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 gender profiles in the areas.

### The AHRQ QI Reference Population

To accomplish risk adjustment, in annual updates of the QI software a **reference population** is analyzed that consists of all HCUP SID data that are available for the year most recently released by AHRQ at the time the QI software is updated. For example when version 4.5 of the QI software was updated in January of 2013 for the May 2013 software release, SID data were available from 2010 from 44 states, so those records serve as the reference population for AHRQ QI software version 4.5.

INSERT text on HCUP data. I believe we have some boilerplate text in some of the other documents – Chris may remember where

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 higher or lower rates than the overall reference population. The reference population rates are published on the AHRQ QI website 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 on the new reference population dataset in an annual process that is described in [Chapter 12](#) of this document. The models are distributed within the QI software, and they facilitate the calculation of risk-adjusted rates. The risk adjustment model covariates and regression coefficients are published on the AHRQ website. See the [links](#) in the Overview chapter of this document.



## Chapter 5. Calculating Area-Level Indicators – Expected, Risk-Adjusted, & Smoothed Rates

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

### The 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 persons living in this specific area. It is predicted for each area using risk-adjustment model coefficients and covariates that summarize the age and gender distribution of the area's population.

### The Risk-Adjusted Rate

The AHRQ QI use indirect standardization to calculate the risk-adjusted rate. 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 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 persons 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 +/- 1.96 \* risk adjusted rate SE] (stored in a data element with a “L” and “U” prefix). (See [Chapter 10](#) section entitled: Computing the Risk-Adjusted Rate Variance. See also [http://qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/Calculating\\_Confidence\\_Intervals\\_for\\_the\\_AHRQ\\_QI.pdf](http://qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/Calculating_Confidence_Intervals_for_the_AHRQ_QI.pdf) )

### The Smoothed Rate

Each area's **smoothed rate** is a weighted average of the risk-adjusted rate and the reference population rate; the smoothed rate is calculated with an empirical Bayes shrinkage estimator to result in a rate near that from the input dataset if the area's rate is estimated in a stable fashion with minimal noise, or 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 a hospital with stable estimates will be similar to the hospital's risk adjusted rate, while the smoothed rate for a hospital with unstable estimates will be more similar to the reference population rate.

The formula for the smoothed rate is:

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

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.

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

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.

For purposes of confidence interval estimation, the *smoothed rate* is assumed to follow a Gamma distribution  $G(shape, scale)$  where

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

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

When there is a fixed comparative rate of interest, it is possible to parameterize the smoothed rate posterior probability based on 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 QI & Present-on-Admission (POA)

**Provider-level indicators** address questions like: 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. They usually include only those cases where 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 where 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 each 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-groups by gender. 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 12](#).

**Present-on-Admission (POA) status** is a third factor that makes provider-level indicators more complex than volume or area-level indicators. Current AHRQ QI that use POA are listed in [Appendix A](#). Some of the indicators look for 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 present on admission, 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 present on admission and should exclude the discharge record from the QI calculation, because the patient is not at risk for the

event. Table 6.1 summarizes those methods, and they are described in more detail in the following chapters, and in [Appendix C](#).

**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 <a href="#">Chapter 7</a> .)	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 can be altered by the User, noting that the User should document any modifications to the program.
2. DX_POA Data Element (See <a href="#">Chapter 8</a> .)	If the diagnosis is flagged as POA using the DX_POA data element, then the record is excluded from the population of interest.	Yes. The user can specify %LET USEPOA = 0; in the CONTROL.SAS program or un-check the WinQI box entitled “Use POA in rate calculation”, either of which will cause the software to ignore DX_POA data that are present in the dataset. Every potential complication will be flagged as an adverse event, and if it does not meet any of the exclusion criteria, it will contribute to the QI numerator. For the purposes of risk-adjustment, a set of coefficients will be employed that were estimated ignoring POA; all complications will be treated as comorbidities.
3. Model the effect of missing data when DX_POA is missing for a particular record, or for the entire dataset (See <a href="#">Chapter 9</a> .)	Use a statistical model included with the QI software and updated annually using reference population data to estimate the probability that the outcome of interest is POA. Use that probability along with the other variables in the record to estimate the probability that the patient experienced the adverse event, conditional on the (possibly large or possibly small) probability that the	Yes. The user can specify %LET USEPOA = 0; in the CONTROL.SAS program or un-check the WinQI box entitled “Use POA in rate calculation”, either of which will cause the software to skip modeling missing POA data.  Alternatively, the user can provide complete POA data, so there is no missing data to be modeled. Note that for

	event was <u>not</u> POA. See <a href="#">Chapter 9</a> and <a href="#">Appendix C</a> .	indicators where POA is a factor in the model, the predicted values are always calculated using the Prediction Module. If the user models the missing POA, then the downstream software uses predictions from the Markov Chain Monte Carlo simulation described in <a href="#">Chapter 9</a> and <a href="#">Appendix C</a> . If the user ignores POA, then the downstream software uses predictions that the Prediction Module calculates using simple scalar multiplication of regression coefficients times covariate values.
--	--	--

## POA Data Element - Background Information

Present-on -Admission 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”<sup>1</sup> or not. This 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 present on admission. 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](http://www.cdc.gov/nchs/data/icd/icd9cm_guidelines_2011.pdf) Again, current AHRQ QI that use POA are listed in [Appendix A](#).

<sup>1</sup> <http://www.cdc.gov/nchs/data/icd9/icdguide10.pdf>.

**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 if condition is present on admission	0	Diagnosis not present at admission
W – Clinically undetermined	Provider is unable to clinically determine whether condition was present on admission or not	1	Diagnosis present at admission
E - Unreported/Not used; Also includes UB-04 values previously coded as "1"	Exempt from POA reporting	1	Diagnosis present at admission

Source: [http://www.cms.hhs.gov/HospitalAcqCond/05\\_Coding.asp#TopOfPage](http://www.cms.hhs.gov/HospitalAcqCond/05_Coding.asp#TopOfPage);  
[http://www.hcup-us.ahrq.gov/db/vars/siddistnote.jsp?var=e\\_poan](http://www.hcup-us.ahrq.gov/db/vars/siddistnote.jsp?var=e_poan).

An individual discharge record might include 20 or more diagnoses. For purposes of the AHRQ QI, the principal diagnosis is always assumed to be present on admission by definition, regardless of the coding of the POA data element in the principal field. Secondary diagnosis codes are considered present on admission if the POA data element is coded with a Y, W, E or 1. Secondary diagnosis codes are considered not present on admission if the POA data element is coded with a N, U or 0.

## Chapter 7. Calculating Provider-Level Observed Rates – Ignoring POA

Provider-level QI calculations are simplest when POA is ignored altogether, so those calculations are described first. Later chapters describe what happens when POA data are present and accounted for, and how the calculations are performed when POA data are missing but modeled. The AHRQ QI software user may ignore the influence of DX\_POA data, either present or missing, by specifying “%LET USEPOA = 0;” in the CONTROL.SAS file or by or un-checking the WinQI box entitled “Use POA in rate calculation”.

When ignoring POA, the main difference between area-level indicators and provider-level indicators is the way the denominator is calculated.

### Discharge Level Indicator Data Element (T)

**Each provider-level observed QI rate consists of a conceptually simple fraction where the denominator is the count of discharge records at risk and the numerator is the count of the 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 is more likely than not to be present on admission and conditions that are POA should not “count” as an adverse event.
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 5 days).
5. Some exclusion criteria are an inherent part of the QI definition (e.g., exclude persons from being at risk for a post-operative hip fracture if the hip repair is the only surgical procedure during the hospitalization).

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

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, 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.

### Consequence of Ignoring POA Data

When POA data are ignored, the observed rate calculation will include records where the outcome of interest was indeed present on admission, and so will inflate the numerator, the denominator, and the observed rate, compared with an **unknown but true underlying rate** that excludes records from population at interest when the outcome was truly POA.



## Chapter 8. Calculating Provider-Level Observed Rates – With Complete POA Data

Consideration of POA 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 will no longer appear erroneously in the numerator, denominator, or observed rate, and the risk adjustment models will no longer 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 depending on the number of records where the outcomes were POA, and with the accuracy of POA coding. This document does not address the topic of POA accuracy. The QI software treats values in the DX\_POA data elements as if they were completely accurate.

The QI calculation procedures are more complicated when some or all of the POA data are missing, so this chapter describes the calculations conducted when POA is present for every record. The following chapter addresses missing POA data.

### Discharge Level POA Exclusion Data Element (Q)

When accounting for POA, the QI software codes the discharge level indicator data element, T, in the same manner described in [Chapter 7](#), using technical specifications to define which records are included in the denominator, numerator, and which should be excluded for one or more reasons. The meaning and possible values of T are described in [Chapter 7](#).

A second, POA-related binary flag is calculated, also. The **discharge level POA exclusion data element** is abbreviated with the letter Q.<sup>2</sup> Put simply, Q records whether the outcome of interest was present on admission or not. The outcome of interest is considered present on admission (Q is assigned “1”) if any of the diagnosis codes that define the outcome of interest are coded as present on admission. 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 (.).”

### The 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

---

<sup>2</sup> The letter P was not available, having been used already for the notion of population at risk. In this 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. (e.g., The variables for PSI 08 are called TPPS08 and QPPS08.)

mean of the T variable after this correction. Note that if POA had been ignored, as in [Chapter 7](#), every record removed from the population at risk by the Q variable would have appeared as a “1” in both the numerator and the denominator. So accounting for POA data yields lower observed rates than when the POA data are ignored. The magnitude of the difference between the rate estimated when POA are ignored and when POA are incorporated will depend on the proportion of records that are flagged as POA that do not meet any of the other indicator exclusion criterion. The accuracy of the difference between the rate estimated when POA are ignored and the rate estimated when POA are incorporated (via the Q flag) depends both on the magnitude of the difference, and the accuracy of the POA coding.

## Chapter 9. Calculating Provider-Level Observed Rates – With Missing POA Data

When POA data are ignored ([Chapter 7](#)) or present in the discharge record ([Chapter 8](#)) then each record in the population at risk contributes a simple “0” or “1” to the QI denominator and if it is a “1” in the denominator, the record contributes a simple “0” or “1” to the numerator. When POA data are missing, the situation is not as simple. Records that do not meet the denominator criteria, regardless of POA, are still simple...they are not in the population at risk. Records where T=0 ignoring POA are simple because they did not experience the outcome of interest, so it could not have been POA. But for other records, the missing DX\_POA flags would determine whether the record was in the population at risk, or not, and if so, whether the patient experienced the outcome of interest. Because we cannot confidently assign a simple 0 or 1 to the numerator and denominator, the QI software calculates expected values of both the numerator and denominator contribution – these expected values fall between 0 and 1, and the software uses them to calculate the observed rate.

The DX\_POA flags can affect the patient record in three ways:

1. The outcome of interest is clearly POA and the record should be excluded from the population at risk.
2. The outcome of interest is clearly not POA and the record should be included in the population at risk.
3. DX\_POA helps distinguish between comorbidities (present at the time of admission) and complications (developed after admission) which affects the assignment of APR-DRG and risk-adjustment.

If some or all of the discharge records in the user’s dataset are missing DX\_POA data elements, the dataset can still be analyzed using methods that take POA into account. The missing POA data are modeled using information from the reference population records that had complete POA data to estimate the expected value of the probability that the outcome of interest was POA, and the expected value of the probability that the patient experienced the outcome of interest if it was not POA.

The expected value calculations use Markov Chain Monte Carlo (MCMC) methods and augmented datasets where the missing POA data are modeled based on relationships observed in the reference population. Specifically, the portion of the reference population dataset where POA was observed yields probabilistic insight into the relative frequency of APR-DRG assignment as well as comorbidities versus complications. In the user’s dataset, if POA data are missing from a discharge record, then the expected values of both the Q flag and the outcome of interest are estimated using an MCMC to approximate the weighted sums over all possible combinations of missing data. The weights in the sums are the probabilities of observing each combination of missing POA flags.

## Prediction Module Nomenclature: $Y = T$ and $P = Q$ and POA improves $Z$ to form $X$

There is a change of nomenclature between the QI software that calculates discharge level data elements and the Prediction Module (PM)<sup>3</sup> software that models the effect of missing POA. In the PM, the outcome is called  $Y$  rather than  $T$  and the POA flag is called  $P$  rather than  $Q$ . In this document we observe this change, and refer to  $Y$  and  $P$  when talking about values that are calculated by the PM.

The set of relevant covariates as coded from the discharge record are collectively described as the vector  $Z$ . After a set of observed or imputed POA flags are applied to the  $Z$  vector, and the covariates are re-calculated, the improved covariates form a vector that we call  $X$ . Data elements that are not affected by POA (e.g., age and gender) take on the same values in the  $Z$  and  $X$  vectors. Data elements that might have changed if POA data had been included with the record (e.g., APR-DRG and comorbidities versus complications) may take different values in  $Z$  and  $X$ . Specifically, the APR-DRG might be changed altogether or shifted to a lower risk of mortality subclass if some of the secondary diagnoses are POA, and some conditions that meet the definition of comorbidity might be changed to complications or vice versa.

## Run Prediction Module to Account for Missing POA

With regard to the observed rate, the Prediction Module does two important things:

1. It calculates the expected value of the probability that the outcome of interest was POA:  $E[P=1 | Y, P, X, Z]$ . Conceptually it does this by imputing POA flags many times and re-calculating whether the outcome of interest was POA. The specifics of the actual MCMC expected value calculation are described in [Appendix C](#). The expected value is a number between 0 and 1 and it is used to determine the record's contribution to the QI denominator. For records with complete POA data, the  $P$  flag takes the value 0 or 1, and the record contributes  $1-P$  to the denominator. (It contributes 1 if the condition is not POA, and 0 if it is POA.) For records with missing POA data, the expected value of  $P$  falls between 0 and 1, and the contribution to the denominator is  $1 - E[P=1 | Y, P, X, Z]$ . That is to say that if there is a 50% chance that the outcome was POA, then the record contributes 0.5 to the denominator. If there is a 99% chance that it was POA, the record contributes 0.01 to the denominator.
2. It calculates the expected value of the contribution of the record to the numerator. If the contribution to the denominator is non-zero and  $Y=1$ , then this probability is equal to  $(1 - \text{the denominator contribution})$ . Otherwise it is zero. If  $T=Y=0$  when POA is missing,

---

<sup>3</sup> The word 'module' is a possible source of confusion. We refer to the four sets of QI as *QI modules: the PQI, IQI, PSI, and PDI*. In a different context, the word 'module' is also used to describe two C++ executable programs that a) estimate the risk-adjustment models during the QI software update process, (the so-called **Analysis Module**) and b) predict the expected value of the outcome in the user's data subject to uncertainty about missing POA (the so-called **Prediction Module**). The Prediction Module is an executable program that is called by SAS or WinQI when the user analyzes their dataset to calculate QI rates.

then Y would not be affected if POA data were present, and the record makes 0 contribution to the QI numerator.

Note that for observations where DX\_POA is present in the user's dataset, the prediction module does not model an expected value. The expected value of Y and P is calculated with certainty as being equal to the observed values T and Q, respectively. These observations make contributions of 1-P (=1-Q) to the denominator, and Y \* (1-P) (which = T \* (1-Q)) to the numerator.

## The Observed Rate

The formula for the observed rate is as follows:

$$\text{Observed Rate} = \frac{\text{Sum of expected value of discharges of the outcome of interest}}{\text{Sum of expected value of discharges in the population at risk}}$$

## Chapter 10. Risk Adjustment for Provider-Level Indicators

This chapter describes risk-adjustment for provider-level QIs. Three special cases are described explicitly: ignoring POA data, accounting for POA in records with complete POA data, and accounting for POA in records with missing POA data.

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. The next section describes the types of data elements that are considered as potential risk-adjusters.

Where possible, the logistic regression models use a generalized estimating equations (GEE) approach to account for correlation at the provider level. When GEE models do not converge during the annual AHRQ QI software update, then multivariable logistic regression models are employed that do not account for that correlation. See [Chapter 12](#) 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 12](#) describes the variable selection process.

For the PSIs, covariates indicate whether the discharge record meets the technical specification for gender, age, modified Diagnosis-Related Group (MDRG) and at least one of twenty-five (25) co-morbidities that are used as covariates in the risk-adjustment model.

For the IQIs, covariates indicate whether the discharge record meets the technical specification for gender, age, All Patient Refined Diagnosis Related Groups (APR-DRG) and risk-of-mortality subclass (minor, moderate, major, extreme) that are used as covariates in the risk-adjustment model.

For the PDIs, covariates indicate whether the discharge record meets the technical specification for birth weight, age in days, age in years, modified Diagnosis-Related Group (MDRG), at least one of forty-six (46) clinical classification software (CCS) co-morbidities and some indicator-specific risk categories that are used as covariates in the risk-adjustment model.

### The Prediction Module

Regardless of whether POA data are ignored or accounted for, and whether the POA data are complete or missing, the provider-level risk adjustment is accomplished using the AHRQ QI Prediction Module software. In the case of accounting for missing POA, it uses an MCMC approach to calculate relevant expected values, as described below. If the user elects to ignore

POA data, or for records where the POA data are complete, then the Prediction Module simply performs scalar multiplication of covariates and coefficients, which is also described below.

## 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. Those files have 21 columns of numbers, and Table 10.1 describes their contents and how they are used in the QI software.

**Table 10.1 Parameter Estimates CSV Files for Provider-Level Risk Adjustment**

Column Number	Column Headings (if any)	Meaning
1-2	Variable Names	List of numbered Z and X covariate names
3-6	[X Z]	Probabilities used for imputation: $P(X=0 Z=0)$ , $P(X=0 Z=1)$ , $P(X=1 Z=0)$ , $P(X=1 Z=1)$ . Note that if these numbers are 1,0,0,1 respectively then the Markov Chain Monte Carlo (MCMC) imputation always imputes $X=Z$ . If the numbers fall between 0 and 1, then sometimes $X = Z$ and sometimes $X \neq Z$ .
7-9	[P X], mse, ese	Regression coefficients (col 7) for the model for POA   X, and their standard errors

Column Number	Column Headings (if any)	Meaning
10-12	[Y Z], mse, ese	Regression coefficients (col 10) for the model for the QI outcome, $Y   Z$ , and their standard errors. This is the model of the outcome that ignores POA. These are the model coefficients that are employed if the user elects to ignore POA. They are multiplied by the Z vector using simple scalar multiplication; the MCMC is not involved in the estimation of this model's parameters or in the computation of the predicted value using this model.
13-15	[Y X], mse, ese	Regression coefficients (col 13) and their standard errors for the model that predicts Y given that POA was coded and therefore the vector X is observed. These coefficients are not used in any calculations that affect the user's output at this time.
16-18	[Y X, P=0], mse, ese	Regression coefficients (col 16) and standard errors for the model that predicts Y given that POA was coded and the outcome Y is known to not have been POA. These coefficients are not used in any calculations that affect the user's output at this time.



Column Number	Column Headings (if any)	Meaning
19-21	[Y X,P=0, MCMC], mse, ese	Regression coefficients (col 19) and standard errors for the model that uses imputed values of POA in the MCMC to predict the probability of the adverse event. These are the coefficients that are typically published in the risk adjustment tables on the AHRQ QI website, and the ones that are used when the user elects to model the effects of missing POA in the data.

## CSV File for the Prediction Module

For each risk-adjusted QI, the software prepares a comma separated values (.csv) file that contains one row per discharge record in the population at risk. The csv file has the following columns:

- **Y** – For purposes of risk-adjustment,  $Y=T$ , the discharge level indicator data element; its value is 0 if the record does not meet the numerator definition, and 1 if it does. Records where T is missing are not at risk for the QI, and are excluded from the QI's csv file.
- **P** – This is the discharge level POA exclusion data element, Q; its value is 0 if  $Y = 0$ ; its value is 0 if  $Y=1$  and the outcome of interest was not POA; its value is 1 if  $Y=1$  and the outcome of interest was POA or met a POA exclusion criterion. If POA is missing, its value is missing.
- **ZCV1 to ZCV $n$**  – A set of  $n$  observed risk-adjustment covariates, each of which is coded using 0/1 indicator data elements. The number of data elements in the vector ( $n$ ), varies from QI to QI. The covariate labels for each QI are listed in the Risk-adjustment Coefficient Tables. (See [links](#) in the Overview.)
- **XCV1 to XCV $n$**  – A vector of  $n$  **enhanced risk-adjustment covariates**. When POA data are present, the vector of XCV values is exactly equal to the vector of ZCV values. When POA data are missing, the XCV values are missing and are modeled in the MCMC.

## Prediction Module Output

The Prediction Module returns a dataset with one row per discharge record, and the following estimated quantities that are used to estimate QI rates. Note that these quantities correspond to the regression models listed in Table 10.1.

- $Y$  – This is the outcome,  $T$ . If POA was observed and the outcome was known to be POA, then  $T$  would have been set to missing and the record would have been eliminated from this dataset.
- $E[Y|Z]$  – This is the expected value of the outcome using the risk-adjustment model that ignores POA data. All comorbidities are treated as POA for the purpose of APR-DRG assignment and for comorbidity terms in the risk-adjustment models. This is the contribution to the numerator of the expected rate if POA is being ignored.
- $E[Y|P=0]$  – This is the contribution to the numerator of the observed rate if POA is being accounted for. If  $Y = 1$  and  $P$  is missing, then this value is  $1 - E[P|YPXZ]$ . If  $P$  is observed, accounted for, and 0, then this is the numerator contribution for the record. If  $P$  is modeled and accounted for, then this is the numerator contribution for the record.
- $E[Y|X, P=0]$ -MCMC – This is the risk-adjusted expected value of the outcome, given  $X$  and  $P=0$ . If POA data are observed, this is simply the scalar product of the risk adjustment coefficients and the risk adjustment covariates. If  $P$  is missing, this quantity is an expected value calculated with the MCMC. This is the contribution to the expected rate numerator if POA is being accounted for. (Note that in the software a small correction is applied to this figure to ensure that the reference population's observed rate equals its expected rate and equals its risk-adjusted rate.)
- $E[P|YPXZ]$  – is the MCMC modeled probability that the outcome was POA. When  $P$  is missing, the denominator contribution of the record is  $1 - E[P|YPXZ]$  and the numerator contribution is between 0 and  $1 - E[P|YPXZ]$ .

## The Expected Rate

To recap, the predicted rate for each discharge comes from the Prediction Module and its method of calculation depends on whether POA is present and being accounted for:

- POA Ignored:  $E[Y|Z]$
- POA Present and accounted for:  $E[Y|X, P=0]$  – MCMC holds the scalar product of the risk adjustment coefficients to calculate  $Y|X, P=0$  and the  $X$  vector.
- POA Absent but accounted for:  $E[Y|X, P=0]$  – MCMC holds the expected value of  $Y$ , calculated by using an MCMC to approximate the weighted average over all possible combinations of missing data.

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

## The Risk-Adjusted Rate

The AHRQ QI use indirect standardization to calculate the risk-adjusted 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 provider or area using a method recommended by Iezzoni and described by Hosmer and Lemeshow that represents the amount of within provider or area variance due to sampling (i.e. as the number of patients per provider or persons 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 +/- 1.96 \* risk adjusted rate SE] (stored in a data element with a “L” and “U” prefix). (See the note below entitled: Computing the Risk-Adjusted Rate Variance. See also [http://qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/Calculating\\_Confidence\\_Intervals\\_for\\_the\\_AHRQ\\_QI.pdf](http://qualityindicators.ahrq.gov/Downloads/Resources/Publications/2011/Calculating_Confidence_Intervals_for_the_AHRQ_QI.pdf) )

## The Smoothed Rate

The formula for the smoothed rate is:

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

where

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

The noise variance is 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.

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

where  $H$  is the number of hospitals with patents 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.

For purposes of confidence interval estimation, the **smoothed rate** is assumed to follow a Gamma distribution  $G(shape, scale)$  where

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

$$scale = \frac{Posterior\ Variance}{Smoothed\ Rate}$$

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

When there is a fixed comparative rate of interest, it is possible to parameterize the smoothed rate posterior probability based on 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

- $E_i$  be the expected (predicted) rate;
- $n_h$  be the number of discharges at hospital  $h$ ; and
- $\alpha$  be the reference population rate (average outcome in the entire sample).

We define the observed rate at hospital  $h$  as

$$O_h = \frac{1}{n_h} \sum_{\substack{i \\ h_i=h}} Y_i$$

the expected rate at hospital  $h$  as

$$E_h = \frac{1}{n_h} \sum_{\substack{i \\ h_i=h}} \hat{\Pi}_i$$

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  $R, S$

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

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  $\hat{\Pi}_i$  (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 non-approximate 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 11](#) on composite measures. Calculation of the smoothed rate is a step in the process of computing the composite measures. However, the basic formula is:

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

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

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

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 based on the Gamma distribution and calculate the probability that the smoothed area rate falls below or above the comparative rate that is of interest.

## Chapter 11. 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 based on the intended use of the composite measure. These weights might be determined empirically or based on non-empirical 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 is first divided by the reference population rate to yield a ratio. The components of the composite are therefore defined in terms of a ratio to the reference population rate for each indicator. The component indicators are scaled by the reference population rate so that each indicator reflects the degree of deviation from the overall average performance.

#### Step 3. Compute the reliability-adjusted ratio

The reliability-adjusted ratio 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).

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

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:

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

*Denominator weight.* A denominator weight is based on the relative frequency of the denominator for each component indicator in the reference population. In general, a denominator weight reflects the 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.

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

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

## 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  $M$  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  $\mu$ , such that:

$$M = \mu + \epsilon \quad (11.1)$$

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

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

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

where  $v$  is a  $1 \times K$  vector with zero mean and  $K \times K$  variance-covariance matrix  $Var(v)$  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  $w$ , this is given by:

$$Var(vw) = w'Var(v)w$$

where  $w'$  indicates the transpose of  $w$ .

Thus, we need an estimate of  $Var(v)$ . We simplify the calculation by assuming that the filtered estimates are formed in isolation for each measure (univariate) and the estimation error is assumed not correlated across measures (e.g., each measure is based on a different sample of patients or independent patient outcomes).

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

$$\hat{\mu}^k = M^k \hat{\beta}^k = M^k (\Omega_\mu^{kk} + \Omega_\epsilon^{kk})^{-1} \Omega_\mu^{kk} \quad (11.3)$$

$$Var(v^k) = \Omega_\mu^{kk} (1 - \hat{\beta}^k) = \Omega_\mu^{kk} - \Omega_\mu^{kk} (\Omega_\mu^{kk} + \Omega_\epsilon^{kk})^{-1} \Omega_\mu^{kk} \quad (11.4)$$

where

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

is the signal ratio of measure  $k$ , the reliability of the measure, and is the r-squared which 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 non-diagonal elements of the covariance matrix (for  $j \neq k$ ):



$$Cov(v^j, v^k) = E[(\mu^j - \hat{\mu}^j)(\mu^k - \hat{\mu}^k)] \quad (11.5)$$

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

$$Cov(v^j, v^k) = \Omega_{\mu}^{jk}(1 - \hat{\beta}^j)(1 - \hat{\beta}^k) \quad (11.6)$$

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

$$lower\ bound = inv\_cdf\_gamma(0.025, shape, scale)$$

$$upper\ bound = inv\_cdf\_gamma(0.975, shape, scale)$$

## Supplemental Notes:

To derive formula (11.6), we substitute

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

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

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

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

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

**QED.**

## Chapter 12. Software Maintenance – Updating the 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 1st, 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<sup>st</sup> of the prior year. In addition, the annual updates include new Census data on the population of counties and new Healthcare Cost and Utilization Project (HCUP) data for the reference population and risk-adjustment covariate coefficients. This chapter describes the methods employed to update the QI reference population and the associated risk-adjustment covariate coefficients.

If the user wishes to account for missing POA, or calculate comparative expected, risk-adjusted, or smoothed rates, then the software makes use of a data frequencies, QI rates, and model coefficients that were estimated using a reference population. In the AHRQ QI software, the reference population consists of all the AHRQ HCUP SID data that are available at the time of the QI update for the year most recently processed. The v4.5 software, released in May 2013, uses 2010 SID data from 44 states for its reference population.

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 (and soon ICD-10) code sets may be modified. Those 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. Those steps are described briefly in [Appendix D](#).

Estimating risk-adjustment models and calculating QI rates in the reference population involves running the QI software on the reference population dataset.

### Assemble the Reference Population Dataset

The user should prepare the input dataset according to the software instructions.

- SID data from all available states are appended together and processed in the manner described in [Chapter 1](#).
- The APR-DRG grouper is run on the adult dataset for the purpose of calculating IQIs. The grouper is run once considering all secondary diagnoses to be POA, and run a second time with POA diagnoses removed. The resulting APR-DRG from the former run is part of the Z vector of IQI covariates and the APR-DRG from the latter run is part of the X vector. This difference captures the fact that when POA is ignored, complications are treated like comorbidities for risk adjustment, and the risk of mortality is probably overstated compared to the risk if the patient were classified using only the conditions that were truly present on admission.

- Missing values of SEX are set to “0” (Male) so they will not be dropped by the QI software. (An alternative would be to impute SEX based on other data elements, like diagnosis codes.)
- Beginning in Version 4.3, discharges from non-community hospitals are deleted from the adult and pediatric analysis data. Community hospitals, as defined by American Hospital Association (AHA), include "all nonfederal, short-term, general and other specialty hospitals, excluding hospital units of institutions." Included among community hospitals are academic medical centers and specialty hospitals such as obstetrics, gynecology, ear nose throat, short-term rehabilitation, orthopedic, and pediatric hospitals. Non-community hospitals include federal hospitals (Veterans Administration, Department of Defense, and Indian Health Service hospitals), long-term hospitals, psychiatric hospitals, alcohol/chemical dependency treatment facilities and hospitals units within institutions such as prisons. (See [http://hcup-us.ahrq.gov/db/state/siddist/siddist\\_hospital.jsp#2008](http://hcup-us.ahrq.gov/db/state/siddist/siddist_hospital.jsp#2008)).
- No other edits are applied to the State Inpatient Databases (SID).

## Calculate Discharge Level Flags

The discharge level T and Q flags are calculated as described in Chapters 3-8.

## Estimate Risk-adjustment Models

There are several steps involved in estimating the QI risk-adjustment models.

1. Construct candidate covariates
2. Select model covariates
3. Estimate the models
4. Evaluate the models

## Construct Candidate Covariates for Risk-adjustment

For the PSIs, potential risk-adjustment covariates indicate whether the discharge record meets the technical specification for gender, age, modified Diagnosis-Related Group (MDRG) and at least one of twenty-five (25) co-morbidities that are used as covariates in the risk-adjustment model.

For the IQIs, potential risk-adjustment covariates indicate whether the discharge record meets the technical specification for gender, age, All Patient Refined Diagnosis Related Groups (APR-DRG) 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, age in days, age in years, modified Diagnosis-Related Group (MDRG), at least one of forty-six (46) clinical classification software (CCS) co-morbidities and some indicator-specific risk categories that are used as covariates in the risk-adjustment model.

For the PQIs, potential risk-adjustment indicate whether the discharge record meets the technical specification for gender, 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 based on 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. For discharge records with POA data, the software creates a second set of data elements (i.e., the **Z** data elements used in the modeling described in [Appendix C](#)) that do not consider secondary diagnosis codes that are not present on admission when assigning comorbidity or risk-of-mortality flags.

## Select Model Covariates

For the provider level indicators, each module has a standard set of covariates grouped into four categories: demographics, severity of illness, comorbidities and other (See [Appendix B](#)). The standard set is tailored to each indicator to create a parsimonious set of covariates for each indicator. Based on 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. For categories that are mutually exclusive, covariates with fewer than 30 cases are pooled into the next covariate along the risk gradient. For example, age 70 to 74 is combined with age 65 to 69, or risk of mortality subclass 3 is combined with subclass 2. For categories with no risk gradient, covariates are pooled into broader covariates. For example, MS-DRGs are pooled into MDCs.

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. The choice of omitted reference category does affect how one might use the model coefficients or odds ratios in an English language sentence, but it does not affect predicted probabilities or model performance.

Once the preliminary multivariable model is specified, it is estimated on the adult or pediatric analytic data, as appropriate. Only those covariates that are statistically significant ( $p < .05$ ) are retained. For covariates that are not statistically significant in categories that are mutually exclusive, the pooling process described above is repeated until a complete, parsimonious model is specified.

For the area level indicators, the models use the complete set of covariates for gender, age in 5-year age groups, an interaction with gender \* 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 website in Risk Adjustment Coefficient Tables. (See [links](#) in the Overview chapter.)

## Estimate the Models

For models where POA is ignored, the AHRQ QI Analysis Module fits a logistic regression model that can be used to calculate the expected value of Y given Z. When possible, the Analysis Module estimates a Generalized Estimating Equations (GEE) model to properly account for within-hospital correlation. If the GEE model does not converge then the Analysis Module fits a more naïve logistic regression model that ignores that extra correlation. Whether the model is a GEE or not may be inferred by the .CSV filename for the QI. For example, PSI 04 uses a file named `gee_pps04_RegressionAnalysisGee.csv`. The ‘Gee’ near the end of the filename indicates that the Analysis Module used a GEE model. On the other hand, PSI 03 uses `gee_pps03_RegressionAnalysis.csv`. The missing ‘Gee’ in the filename tells the user that the model is not a GEE.

When POA data is accounted for, the Analysis Module uses Markov Chain Monte Carlo (MCMC) methods to fit several models.

1. It estimates coefficients to predict the expected value of Y given X, P=0 for records where POA is observed.
2. It estimates coefficients for a model for the expected value of the discharge level POA exclusion data element (P) when POA is missing.
3. And it estimates coefficients for the Prediction Module to calculate the expected value of the outcome, Y given P=0 and the observed data, for missing POA.

Computational details are described in [Appendix C](#). The Analysis Module generates a comma-separated values (.csv) file for each risk-adjusted QI that the Prediction Module uses when applying the models to a user’s dataset. These files are part of the AHRQ QI software package that is made available on the AHRQ website. See Table 10.1 for a description of the contents of those .csv files.

## Calculate Rates

After the new risk-adjustment models are fit, the Prediction Module is run on the data to calculate expected values for P and Y so that observed rates may be calculated for the reference population. Reference population rates and signal variances are calculated both ignoring POA altogether and accounting for missing POA. 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.

## Update Software

In addition to the aforementioned .CSV and .TXT files, the AHRQ QI software must be updated to generate and combine the correct set of covariate variables for each risk adjusted QI. These covariates are generated in the so-called ~SAS3.SAS programs, and whenever the list of covariates in a risk-adjustment model changes, that code must be changed accordingly. Note that

it possible to fit new risk-adjustment model coefficients without updating the list of covariates. In that case, the ~SAS3.SAS program may need very little revision, if any.

## 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 be 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 where 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 where 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 website in the Risk Adjustment Coefficient Tables. (See [links](#) in the Overview chapter.)

### Calibration

Calibration is often described by sorting the dataset based on 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 logistic regression goodness-of-fit statistic (1980) is based on a chi-square test statistic calculated using the observed and expected counts across the ten 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 ten deciles of risk.

## Chapter 13. Software Maintenance – Other Updates

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 metropolitan statistical area is crucial for calculating the denominators of the area-level measures. That file and the method to construct it are described in a file entitled *AHRQ QI Population File Documentation* on the AHRQ website: (<http://www.qualityindicators.ahrq.gov/software/SAS.aspx>)

### Condition-Specific Population File

The AHRQ QI program is conducting current methods 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 QICTYC13.TXT that is included with the v4.5 AHRQ QI module. That file was calculated using the following steps:

1. Use the population reference file to estimate 2013 population for each combination of state and age category. In the QI software, age categories are coded as:

```
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://www.statehealthfacts.org/comparebar.jsp?ind=73&cat=2>.
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 2013. (Population data from 2013 and proportion data from 2010.)



## References

- Dempster AP; Laird NM; Rubin DB (1977). Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society, Series B (Methodological)* 39 (1): 1–38.
- Dimick JB, Staiger DO, Birkmeyer JD. (2006) Are Mortality Rates for Different Operations Related? Implications for measuring the quality of noncardiac surgery. *Med Care* Aug;44(8):774-8
- Fitzmaurice GM, Laird NM and Ware JH. *Applied Longitudinal Analysis*. John Wiley & Sons, Inc., Thirteenth edition, 2004.
- Hammersley JM & Handscomb DC, (1964) *Monte Carlo Methods*. Methuen & Co Ltd, London.
- Hosmer DW & Lemeshow S. (2000) *Applied Logistic Regression*, 2<sup>nd</sup> ed., Wiley, Inc., New York.
- Hosmer DW & Lemeshow S. (1980) A goodness-of-fit test for the multiple logistic regression model. *Communications in Statistics*, A10, 1043-1069.
- Hosmer DW & Lemeshow S. (1995). Confidence interval estimates of an index of quality performance based on logistic regression. *Statistics in Medicine*, Vol 14, Issue 19, 2161-2172
- Iezzoni, Lisa, ed. (2013) *Risk Adjustment for Measuring Health Care Outcomes*, 4<sup>th</sup> ed. Health Administration Press, Chicago.
- Liang KY & Zeger SL. Longitudinal Data Analysis Using Generalized Linear Models. *Biometrika*, 73(1):13-22, April 1986
- Little RJA & Rubin DB (2002). *Statistical Analysis with Missing Data*. Wiley, Hoboken, NJ.
- Luft, HS & Brown, BW Jr. (1993). Calculating the Probability of Rare Events: Why Settle for an Approximation? *Health Services Research* 28:4, 419-439
- McClellan M & Staiger D. (1999) The quality of health care providers. Cambridge, MA: National Bureau of Economic Research. NBER Working Paper #7327. Available at: <http://www.nber.org/papers/w7327>.
- Morris, CN (1983) Parametric Empirical Bayes Inference: Theory and Applications, *Journal of the American Statistical Association*, Vol. 78, No. 381 (Mar., 1983), pp. 47-55.
- Press WH., Teukolsky SA, Vetterling WT, Flannery BP, (1992) *Numerical recipes in C (2nd ed.): the art of scientific computing*. Cambridge University Press New York, NY, USA.
- Robert CP & Casella G. (2004). *Monte Carlo Statistical Methods*, 2<sup>nd</sup> Ed. Springer, New York, NY.
- Ripley BD, (1987) *Stochastic Simulation*. New York: Wiley & Sons, NY
- Rubinstein BY, (1981) *Simulation and the Monte Carlo Method*. Wiley & Sons., NY

Zeger SL & Liang KY. Longitudinal Data Analysis for Discrete and Continuous Outcomes.  
*Biometrics*, 42(1):121-130, March 1986.

## Appendix A. Table of AHRQ QI Risk-adjustment / POA

Appendix Table A.1 denotes which AHRQ QI are risk-adjusted and which use POA data and for what purpose (i.e., for technical specifications or risk-adjustment).

An entry of ‘AM/PM’ in the column entitled ‘Calculate Risk Adjusted Rate’ means that the indicator is a provider-level indicator and its risk adjustment model is estimated using the Analysis Module (AM) described in [Appendix C](#). The risk adjustment calculations are carried out using the Prediction Module (PM), also described in [Appendix C](#). An entry of ‘SAS’ in the column entitled ‘Calculate Risk Adjusted Rate’ means that the indicator is an area-level indicator and its risk adjustment model is estimated using PROC LOGISTIC in SAS.

An X in the column marked ‘Technical Specifications’ means that the indicator has an exclusion that explicitly references the POA data element. A QI software user may tell the software to ignore the DX\_POA data element for purposes of risk-adjustment, but the software will never ignore DX\_POA if it is referenced in the technical specifications for the purpose of defining exclusions, and if the data element is present in the discharge record. When a discharge record is missing the DX\_POA data element, the Q flag will be set to “missing (.)” and the software will either ignore it (if USEPOA=0) or impute it (if USEPOA=1).

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 covariates as comorbidities if the user elects to ignore POA data
- Model the missing POA data via the Markov Chain Monte Carlo (MCMC) if the user elects to account for POA data.

See [Chapter 10](#) for additional details on risk adjustment.

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

	Calculate Risk Adjusted Rate	Use POA?	
		Technical Specifications	Risk Adjustment
IQI 01 - Esophageal Resection Volume			
IQI 02 - Pancreatic Resection Volume			
IQI 04 - Abdominal Aortic Aneurysm (AAA) Repair Volume			

	Calculate Risk Adjusted Rate	Use POA?	
		Technical Specifications	Risk Adjustment
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	AM/PM		X
IQI 09 - Pancreatic Resection Mortality Rate	AM/PM		X
IQI 11 - AAA Repair Mortality Rate	AM/PM		X
IQI 12 - CABG Mortality Rate	AM/PM		X
IQI 13 - Craniotomy Mortality Rate	AM/PM		X
IQI 14 - Hip Replacement Mortality Rate	AM/PM		X
IQI 15 - Acute Myocardial Infarction (AMI) Mortality Rate	AM/PM		X
IQI 16 - Heart Failure Mortality Rate	AM/PM		X
IQI 17 - Acute Stroke Mortality Rate	AM/PM		X
IQI 18 - Gastrointestinal Hemorrhage Mortality Rate	AM/PM		X
IQI 19 - Hip Fracture Mortality Rate	AM/PM		X
IQI 20 - Pneumonia Mortality Rate	AM/PM		X
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	SAS		
IQI 27 - Percutaneous Coronary Intervention (PCI) Rate	SAS		
IQI 28 - Hysterectomy Rate	SAS		
IQI 29 - Laminectomy or Spinal Fusion Rate	SAS		
IQI 30 - Percutaneous Coronary Intervention (PCI) Mortality Rate	AM/PM		X
IQI 31 - Carotid Endarterectomy Mortality Rate	AM/PM		X

	Calculate Risk Adjusted Rate	Use POA?	
		Technical Specifications	Risk Adjustment
IQI 32 - Acute Myocardial Infarction (AMI) Mortality Rate, Without Transfer Cases	AM/PM		X
IQI 33 - Primary Cesarean Delivery Rate, Uncomplicated			
IQI 34 - Vaginal Birth After Cesarean (VBAC) Rate, All			
PSI 02 - Death Rate in Low-Mortality Diagnosis Related Groups (DRGs)	AM/PM		X
PSI 03 - Pressure Ulcer Rate	AM/PM	X	X
PSI 04 - Death Rate among Surgical Inpatients with Serious Treatable Complications	AM/PM		X
PSI 05 - Retained Surgical Item or Unretrieved Device Fragment Count		X	
PSI 06 - Iatrogenic Pneumothorax Rate	AM/PM	X	X
PSI 07 - Central Venous Catheter-Related Blood Stream Infection Rate	AM/PM	X	X
PSI 08 - Postoperative Hip Fracture Rate	AM/PM	X	X
PSI 09 - Perioperative Hemorrhage or Hematoma Rate	AM/PM	X	X
PSI 10 - Postoperative Physiologic and Metabolic Derangement Rate	AM/PM	X	X
PSI 11 - Postoperative Respiratory Failure Rate	AM/PM	X	X
PSI 12 - Perioperative Pulmonary Embolism or Deep Vein Thrombosis Rate	AM/PM	X	X
PSI 13 - Postoperative Sepsis Rate	AM/PM	X	X
PSI 14 - Postoperative Wound Dehiscence Rate	AM/PM		X
PSI 15 - Accidental Puncture or Laceration Rate	AM/PM	X	X
PSI 16 - Transfusion Reaction Count		X	
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			
PDI 01 - Accidental Puncture or Laceration Rate	AM/PM	X	X
PDI 02 - Pressure Ulcer Rate	AM/PM	X	X
PDI 03 - Retained Surgical Item or Unretrieved Device Fragment Count		X	
PDI 05 - Iatrogenic Pneumothorax Rate	AM/PM	X	X

	Calculate Risk Adjusted Rate	Use POA?	
		Technical Specifications	Risk Adjustment
PDI 06 - RACHS-1 Pediatric Heart Surgery Mortality Rate	AM/PM		X
PDI 07 - RACHS-1 Pediatric Heart Surgery Volume			
PDI 08 - Perioperative Hemorrhage or Hematoma Rate	AM/PM	X	X
PDI 09 - Postoperative Respiratory Failure Rate	AM/PM	X	X
PDI 10 - Postoperative Sepsis Rate	AM/PM	X	X
PDI 11 - Postoperative Wound Dehiscence Rate			X
PDI 12 - Central Venous Catheter-Related Blood Stream Infection Rate	AM/PM	X	X
PDI 13 - Transfusion Reaction Count		X	
PDI 14 – Asthma Admission Rate	SAS		
PDI 15 – Diabetes Short-Term Complications Admission Rate	SAS		
PDI 16 – Gastroenteritis Admission Rate	SAS		
PDI 17 – Perforated Appendix Admission Rate	SAS		
PDI 18 – Urinary Tract Infection Admission Rate	SAS		
NQI 01 - Neonatal Iatrogenic Pneumothorax Rate		X	X
NQI 02 - Neonatal Mortality Rate	AM/PM		X
NQI 03 - Neonatal Blood Stream Infection Rate	AM/PM	X	X
PQI 01 - Diabetes Short-Term Complications Admission Rate	SAS		
PQI 02 - Perforated Appendix Admission Rate	SAS		
PQI 03 - Diabetes Long-Term Complications Admission Rate	SAS		
PQI 05 - Chronic Obstructive Pulmonary Disease (COPD) or Asthma in Older Adults Admission Rate	SAS		
PQI 07 - Hypertension Admission Rate	SAS		
PQI 08 - Heart Failure Admission Rate	SAS		
PQI 09 - Low Birth Weight Rate	SAS		
PQI 10 - Dehydration Admission Rate	SAS		
PQI 11 - Bacterial Pneumonia Admission Rate	SAS		

	Calculate Risk Adjusted Rate	Use POA?	
		Technical Specifications	Risk Adjustment
PQI 12 - Urinary Tract Infection Admission Rate	SAS		
PQI 13 - Angina Without Procedure Admission Rate	SAS		
PQI 14 - Uncontrolled Diabetes Admission Rate	SAS		
PQI 15 - Asthma in Younger Adults Admission Rate	SAS		
PQI 16 - Lower-Extremity Amputation Among Patients With Diabetes Rate	SAS		

IQI = Inpatient Quality Indicator; PSI = Patient Safety Indicator; PDI = Pediatric Quality Indicator; NQI = Neonatal Quality Indicator

## 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, age 18-24 with age 25-29 or combine ROM subclass “4” with ROM subclass “3”.



**Appendix Table B.1 Table of 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 to 1 year)  Age in years (1 year and above)	Birth weight (500g groups)
Severity of Illness	DRGs pool into MDCs	APR-DRG  Major Diagnosis Categories (MDC)	Modified MS-DRG*  Major Diagnosis Categories (MDC)	Modified MS-DRG*  Major Diagnosis Categories (MDC)	Modified MS-DRG*  Major Diagnosis Categories (MDC)
Comorbidities		APR-DRG  Risk of mortality subclass  (1 – minor; 2 - moderate; 3 – major; 4 – extreme)	AHRQ Comorbidities	AHRQ Clinical Classification Software	Congenital anomalies
Other		Transfer-in status	Transfer-in status	Transfer-in status	Transfer-in status
		Point of Origin status	Point of Origin status  Days to Procedure status	Point of Origin status  Days to Procedure status  Indicator-specific risk stratifiers	Point of Origin status  Days to Procedure status

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

## Appendix C. Provider-Level Risk Adjustment - Detailed Methods

This appendix gives some statistical detail about how the provider-level risk adjustment models are fit and how they account for missing POA data. The Analysis Module is described first. It is used annually to fit models that are incorporated into updated AHRQ QI software. The Prediction Module is described second. It is called by the SASP3.SAS program for the IQI, PSI, and PDI indicators.

### The Analysis Module

The purpose of the Analysis Module (AM) is to fit a set of regression coefficients using the data of the reference population. The input dataset is expected to have variables corresponding to the outcome of interest at discharge  $Y$ , one or more indicators of an outcome of interest present on admission (POA indicators  $P$ ), and covariate vectors  $\mathbf{X}$  and  $\mathbf{Z}$  containing demographic, condition, co-morbidity, and potentially any other information, used as explanatory variables. The covariate  $\mathbf{X}$  is considered an improved measurement of the quantities measured by the covariate  $\mathbf{Z}$ .

Conceptually, there could be many ways in which  $\mathbf{X}$  might improve  $\mathbf{Z}$ . At this time, those improvements are the following:

1. In the  $\mathbf{Z}$  vector, the discharge level POA exclusion data element is sometimes observed and sometimes missing. In the  $\mathbf{X}$  vector, the missing values are modeled, when missing.
2. The  $\mathbf{Z}$  vector uses all secondary diagnoses to assign APR-DRG for the IQI, but the  $\mathbf{X}$  vector uses the DX\_POA data element (observed or modeled) to take into account only the diagnoses that were present on admission.
3. The  $\mathbf{Z}$  vector considers all comorbidities to be complications for the purpose of calculating the observed rate and all to be comorbidities for the purpose of risk-adjustment. The  $\mathbf{X}$  vector uses the DX\_POA data element (observed or modeled) to distinguish between complications and comorbidities.

The outcome  $Y$  and covariate  $\mathbf{Z}$  variables are never missing, but elements of the covariate  $\mathbf{X}$  and values of the present-on-admission indicators  $P$  can be missing. The dataset also contains a hospital identification number and a record identification number (a key identifying unique discharge records.)

### Missing Data

Missing data are handled by integrating the likelihood over all the possible values of the missing variables. This technique for dealing with missing data is well-established in the statistical literature. Little and Rubin (2002) devote several chapters to analyzing missing data by integrating over the distribution, or likelihood, of the missing data. When the integral (or sum) of

the likelihood cannot be feasibly calculated, an alternative method known as the Expectation-Maximization (EM) algorithm can be used. The EM algorithm was developed in the 1970s by Dempster, Laird and Rubin (1977) to solve MLE equations in the presence of missing data. More recently, related methods based on Markov chain Monte Carlo (MCMC) algorithms have become popular for dealing with missing and censored data. MCMC algorithms include methods such as Metropolis-Hastings or Gibbs sampling which are widely used in Bayesian statistical analysis (Robert and Casella, 2004). MCMC methods are general and robust, and can be applied to a large variety of models. These methods are based on simulation, and they produce results that are approximations of the value being estimated. The approximation error can be controlled by the number of MCMC steps used in the simulation. In particular, as the number of MCMC steps goes to infinity, the approximation error goes to zero. We will give detail about the MCMC used in the Analysis and Prediction Modules in the following sections.

## Data Notation

Here is the general statistical notation used to describe the model:

- $h_i$  is the hospital associated with the  $i^{th}$  record (patient);
- $Y_i$  is a binary variable indicating the outcome of interest at hospital discharge associated with the  $i^{th}$  record.  $Y_i = 1$  if the patient experiences the outcome of interest,  $Y_i = 0$  otherwise;
- $P_i$  is a binary variable indicating whether an outcome of interest is present on admission. Notice that if  $Y_i = 0$ , then it is assumed that  $P_i = 0$ . If more than one POA indicators are present, the maximum value is considered;
- $\mathbf{Z}_i$  is a vector of binary explanatory variables associated with the  $i^{th}$  record;
- $\mathbf{X}_i$  is a vector of improved binary explanatory variables associated with the  $i^{th}$  record.

In the following formulae  $i$  indicates the record index while  $k$  indicates the component index of the covariate vectors. For example, indicating with  $K$  the number of components of the covariate vectors, then  $\mathbf{X}_i \in R^K$  indicates the vector of covariates associated with the  $i^{th}$  record,  $X_{ik}$  indicates the value of the  $k^{th}$  covariate associated with the  $i^{th}$  record, while  $X_k$  without the record index is used to indicate the  $k^{th}$  covariate of a generic covariate vector.

The description of the Analysis Module proceeds with a brief outline of the MCMC calculations to account for missing POA data. The outline is a conceptual simplified description using formulae without explicit posterior parameters. The sections after the outline give additional detail

## Outline of the MCMC algorithm to fit $[Y|X, P = 0]$ on data sets with missing data.

Before MCMC begins:

- Fit 2 by 2 binary tables  $[X_k|Z_k]$  using observations where both  $X_k$  and  $Z_k$  are measured;
- Fit logistic regression model for  $[P|X]$  using observations with complete data.

MCMC loop:

1. Build joint distribution  $[Y, X, P, Z] = [Y|X, P][P|X][X|Z]$ ;
2. Use full conditional distribution  $[X|Y, P, Z] = [Y, X, P, Z]/[Y, P, Z]$  to draw missing  $X$ s (Gibbs Sampling); write the drawn missing  $X$  values to the chain. This values can be referred to as *imputed data*;
3. Use full conditional distribution  $[P|Y, X, Z] = [Y, X, P, Z]/[Y, X, Z]$  to draw missing  $P$ s (Gibbs Sampling); write the drawn missing  $P$  values to the chain. This values can be referred to as *imputed data*;
4. Fit logistic model  $[Y|X, P = 0]$  using the available data, where measured, and the last imputed data, where  $X$  and  $P$  are missing (see 2. and 3.) Use either MLE or GEE (depending on the user's choice) to fit the model and obtain the estimated  $\hat{\beta}$  and the estimated  $\text{var}(\hat{\beta})$ ;
5. Draw a new set of regression coefficients  $\beta$  from a multivariate normal distribution with mean  $\hat{\beta}$  and variance  $\text{var}(\hat{\beta})$ ; write the drawn  $\beta$  values to the chain.
6. Go to 1 until total iterations equals that specified in the input XML file.

Note: the probability distribution density  $[Y|X, P]$  is equal to the model probability distribution density  $[Y|X, P = 0]$  when  $P = 0$ ; and it is equal to the marginal probability  $[Y = 1]$  when  $P = 1$ .

During the MCMC loop:

- Drop burn-in entries, as specified in input XML file.
- Thin the chain, as specified in input XML file.

After the loop:

- Estimate the regression coefficients  $\hat{\beta}_{MCMC}$  and their standard error by calculating the expected values (mean) and the standard deviation of the components of the MCMC chain representing the regression coefficients  $\beta$  associated with the model  $[Y|X, P = 0]$ .

## More Detailed Statistical Model

The main goal of the model is the estimation of  $Y$  given  $\mathbf{X}$  and  $P = 0$ . We assume the “conditional” binomial model

$$[Y|\mathbf{X}, P; \beta_Y] = \prod_i (\pi_{Y,i}^{1-P_i})^{Y_i} (1 - \pi_{Y,i}^{1-P_i})^{1-Y_i} \quad (\text{C.1})$$

with logistic link

$$\text{logit}(\pi_{Y,i}) = X_i \beta_Y$$

Another component of the model is the estimation of  $P$  given  $\mathbf{X}$ , which is used to predict  $P$  when that value is missing. We assume the binomial model

$$[P|\mathbf{X}; \boldsymbol{\beta}_P] = \prod_i \pi_{P,i}^{Y_i} (1 - \pi_{P,i})^{1-Y_i} \quad (\text{C.2})$$

with logistic link

$$\text{logit}(\pi_{P,i}) = X_i \boldsymbol{\beta}_P$$

Furthermore, we estimate  $\mathbf{X}$  when elements of that vector are missing by using the information contained in  $\mathbf{Z}$ . Since both  $\mathbf{X}$  and  $\mathbf{Z}$  contain binary variables, we model  $[\mathbf{X}|\mathbf{Z}]$  using the two vectors of probabilities

$$\pi_{X,k}(0) = \Pr[X_k = 1|Z_k = 0]$$

$$\pi_{X,k}(1) = \Pr[X_k = 1|Z_k = 1]$$

and the likelihood

$$[\mathbf{X}|\mathbf{Z}; \boldsymbol{\pi}_X] = \prod_{ik} \pi_{X,ik}^{X_{ik}} (1 - \pi_{X,ik})^{1-X_{ik}} \quad (\text{C.3})$$

where

$$\pi_{X,ik} = \pi_{X,k}(Z_{ik})$$

Combining equations (C.1), (C.2) and (C.3), we obtain the likelihood

$$\begin{aligned} L(Y, \mathbf{X}, P, \mathbf{Z}; \boldsymbol{\beta}_Y, \boldsymbol{\beta}_P, \boldsymbol{\pi}_X) &= [Y, \mathbf{X}, P|\mathbf{Z}; \boldsymbol{\beta}_Y, \boldsymbol{\beta}_P, \boldsymbol{\pi}_X] = \\ &= [Y|\mathbf{X}, P; \boldsymbol{\beta}_Y] \times [P|\mathbf{X}; \boldsymbol{\beta}_P] \times [\mathbf{X}|\mathbf{Z}; \boldsymbol{\pi}_X] = \\ &= \prod_i (\pi_{Y,i}^{1-P_i})^{Y_i} (1 - \pi_{Y,i}^{1-P_i})^{1-Y_i} \pi_{P,i}^{P_i} (1 - \pi_{P,i})^{1-P_i} \pi_{X,ik}^{X_{ik}} (1 - \pi_{X,ik})^{1-X_{ik}} \end{aligned} \quad (\text{C.4})$$

Likelihood (C.4) is written as a distribution of  $Y, \mathbf{X}, P$  given  $\mathbf{Z}$ . In order to write the model for missing  $\mathbf{X}$  and  $P$ , we introduce the “true” variables  $\mathbf{X}', P'$  (to which we refer as “imputed”) and add the data model

$$[X'_{ik}|X_{ik}] = \begin{cases} X_{ik} & X_{ik} \text{ is measured} \\ 1/2 & \text{otherwise} \end{cases} \quad (\text{C.5})$$

$$[P'_i|P_i] = \begin{cases} P_i & P_i \text{ is measured} \\ 1/2 & \text{otherwise} \end{cases} \quad (\text{C.6})$$

The data model acts as a family of indicator variables, fixing the “imputed” variable to the measured value if the data are not missing. The likelihood integrated (summed) over the missing data can now be written as

$$\tilde{L}(Y, \mathbf{X}, P, \mathbf{Z}; \boldsymbol{\beta}_Y, \boldsymbol{\beta}_P, \boldsymbol{\pi}_X) = \sum_{P', \mathbf{X}'} L(Y, \mathbf{X}', P', \mathbf{Z}; \boldsymbol{\beta}_Y, \boldsymbol{\beta}_P, \boldsymbol{\pi}_X) \times [\mathbf{X}'|\mathbf{X}] \times [P'|P] =$$

$$= \sum_{P', X'} [Y|X', P'; \beta_Y] \times [P'|X'; \beta_P] \times [X'|Z; \pi_X] \times [X'|X] \times [P'|P] \quad (C.7)$$

Since the distribution inside the sum is the product of distributions for each record  $i$ , (see equation C.4), using the distributive property we can write

$$\begin{aligned} \tilde{L}(Y, X, P, Z; \beta_Y, \beta_P, \pi_X) &= \\ &= \prod_i \left\{ \sum_{P'_i, X'_i} [Y_i|X'_i, P'_i; \beta_Y] \times [P'_i|X'_i; \beta_P] \times [X'_i|Z_i; \pi_X] \times [X'_i|X_i] \times [P'_i|P_i] \right\} \end{aligned}$$

As the number of components of the covariate vector  $X$  increases, to compute the above sum deterministically becomes unfeasible. For example, if  $X_i$  has 30 components, then the number of sums for every record  $i$  with missing  $X_i$  data is  $2^{30} > 10^9$ , and if the number of components is 100, then the number of sums becomes  $2^{100} > 10^{30}$ . The AM and PM employ alternative methods for integrating (summing) the likelihood over the missing data.

## Model Fitting Approach using MCMC

To fit the  $\beta_Y$  coefficients using the *marginal* likelihood (C.7) (that is, the likelihood integrated over the missing data), we use Gibbs sampling, which is a standard MCMC technique (see Robert and Casella, 2004).

After reading the data, the AM fits the coefficients  $\hat{\beta}_P$  and  $\hat{\pi}_X$  using only the records in the dataset that have no missing data. Then, given  $\hat{\beta}_P$  and  $\hat{\pi}_X$ , a sample of values of  $\beta_Y$ ,  $X'$ , and  $P'$  is drawn from the posterior distribution:

$$[X', P', \beta_Y]_{post} \propto [Y|X', P'; \beta_Y] \times [P'|X'; \hat{\beta}_P] \times [X'|Z; \hat{\pi}_X] \times [X'|X] \times [P'|P] \quad (C.8)$$

The posterior distribution factors as

$$[X', P', \beta_Y]_{post} = \prod_i [X'_i, P'_i, \beta_Y]_{post}$$

Univariate and multivariate Gibbs sampling is used to sample  $X'$ ,  $P'$ , and  $\beta_Y$ . The sampling equations are the following:

- Sampling of  $P'_i$  (univariate Gibbs sampling)

$$P'_{i,new} \sim [P'_i|X'_i, \beta_Y]_{post} = \frac{[X'_i, P'_i, \beta_Y]_{post}}{[X'_i, \beta_Y]_{post}} = \frac{[X'_i, P'_i, \beta_Y]_{post}}{[X'_i, P'_i = 0, \beta_Y]_{post} + [X'_i, P'_i = 1, \beta_Y]_{post}}$$

Notice that posterior conditional distribution  $[P'_i | \mathbf{X}'_i, \boldsymbol{\beta}_Y]_{post}$  is defined from the joint posterior on the left-hand-side of (C.8), and it is different from the conditional distribution  $[P' | \mathbf{X}'; \hat{\boldsymbol{\beta}}_P]$ , which appears on the right-hand-side of (C.8). Due to the constraint  $Y_i = 0 \Rightarrow P_i = 0$ , we have  $P'_{i,new} = 0$  if  $Y_i = 0$ . When  $Y_i = 1$ , using equation (C.8) and simplifying the common factors in the numerator and the denominator, we can write

$$\begin{aligned} [P'_i = 1 | \mathbf{X}'_i, \boldsymbol{\beta}_Y; Y_i = 1]_{post} &= \\ &= \frac{[Y_i = 1 | \mathbf{X}'_i, P'_i = 1; \boldsymbol{\beta}_Y] \times \hat{\pi}_{P,i}}{[Y_i = 1 | \mathbf{X}'_i, P'_i = 0; \boldsymbol{\beta}_Y] \times (1 - \hat{\pi}_{P,i}) + [Y_i = 1 | \mathbf{X}'_i, P'_i = 1; \boldsymbol{\beta}_Y] \times \hat{\pi}_{P,i}} \end{aligned}$$

where  $\hat{\pi}_{P,i}$  is the estimated probability<sup>4</sup>  $[P'_i = 1 | \mathbf{X}'_i; \hat{\boldsymbol{\beta}}_P]$ . Noticing that  $[Y_i = 1 | \mathbf{X}'_i, P'_i = 1; \boldsymbol{\beta}_Y] = 1$  and  $[Y_i = 1 | \mathbf{X}'_i, P'_i = 0; \boldsymbol{\beta}_Y] = \pi_{Y,i}$ , we obtain

$$[P'_i = 1 | \mathbf{X}'_i, \boldsymbol{\beta}_Y; Y_i = 1]_{post} = \frac{\hat{\pi}_{P,i}}{\pi_{Y,i}(1 - \hat{\pi}_{P,i}) + \hat{\pi}_{P,i}}$$

Hence, the sampling equations for  $P'_{i,new}$  become

$$\begin{aligned} P'_{i,new} &= 0, & \text{if } Y_i = 0 \\ P'_{i,new} &\sim \text{Bernoulli}\left(\frac{\hat{\pi}_{P,i}}{\pi_{Y,i}(1 - \hat{\pi}_{P,i}) + \hat{\pi}_{P,i}}\right), & \text{if } Y_i = 1 \end{aligned}$$

- Sampling of  $X'_{ik}$  (univariate Gibbs sampling)

$$\begin{aligned} X'_{ik,new} \sim [X'_{ik} | \mathbf{X}'_{ik-}, P'_i, \boldsymbol{\beta}_Y]_{post} &= \frac{[X'_{ik}, \mathbf{X}'_{ik-}, P'_i, \boldsymbol{\beta}_Y]_{post}}{[\mathbf{X}'_{ik-}, P'_i, \boldsymbol{\beta}_Y]_{post}} = \\ &= \frac{[X'_{ik}, \mathbf{X}'_{ik-}, P'_i, \boldsymbol{\beta}_Y]_{post}}{[X'_{ik} = 0, \mathbf{X}'_{ik-}, P'_i, \boldsymbol{\beta}_Y]_{post} + [X'_{ik} = 1, \mathbf{X}'_{ik-}, P'_i, \boldsymbol{\beta}_Y]_{post}} \end{aligned}$$

where  $\mathbf{X}'_{ik-}$  indicates all the components of the vector  $\mathbf{X}'_i$  except the  $k$ -th one, and  $[X'_{ik}, \mathbf{X}'_{ik-}, P'_i, \boldsymbol{\beta}_Y]_{post} \equiv [\mathbf{X}'_i, P'_i, \boldsymbol{\beta}_Y]_{post}$ . Using equation (C.8) and simplifying the common factors in the numerator and the denominator, we can write

$$[X'_{ik} = 1 | \mathbf{X}'_{ik-}, P'_i, \boldsymbol{\beta}_Y]_{post} = \frac{f(X'_{ik} = 1) \times \hat{\pi}_{X,ik}}{f(X'_{ik} = 0) \times (1 - \hat{\pi}_{X,ik}) + f(X'_{ik} = 1) \times \hat{\pi}_{X,ik}}$$

where

<sup>4</sup> As mentioned at the beginning of the paragraph, the Analysis Module estimates  $\hat{\boldsymbol{\beta}}_P$  and  $\hat{\pi}_X$  using only the records with no-missing data before the MCMC analysis.

$$f(X'_{ik}) = [Y_i | X'_{ik}, \mathbf{X}'_{ik-}, P'_i; \boldsymbol{\beta}_Y] \times [P'_i | X'_{ik}, \mathbf{X}'_{ik-}; \hat{\boldsymbol{\beta}}_P]$$

and  $\hat{\pi}_{X,ik}$  is the estimated probability<sup>4</sup>  $[X_{ik} = 1 | Z_{ik}]$ . Hence, the sampling equation for  $X'_{ik,new}$  become

$$X'_{ik,new} \sim \text{Bernoulli} \left( \frac{f(X'_{ik} = 1) \times \hat{\pi}_{X,ik}}{f(X'_{ik} = 0) \times (1 - \hat{\pi}_{X,ik}) + f(X'_{ik} = 1) \times \hat{\pi}_{X,ik}} \right)$$

- Sampling of  $\boldsymbol{\beta}_Y$  (multivariate Gibbs sampling)

$$\boldsymbol{\beta}_{Y,new} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \times N(\mathbf{0}, \sigma^2 \mathbf{I})$$

where  $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$  is the multivariate normal approximation of the function

$$\boldsymbol{\beta}_Y \rightarrow [Y | \mathbf{X}', P'; \boldsymbol{\beta}_Y] = \prod_i \left( \pi_{Y,i}^{1-P'_i} \right)^{Y_i} \left( 1 - \pi_{Y,i}^{1-P'_i} \right)^{1-Y_i}$$

using a second order Taylor expansion of the log-likelihood, as standard practice in Generalized Linear Models.

The AM includes an option to use Generalized Estimating Equations (Zeger & Liang, 1986, Liang & Zeger, 1986, Fitzmaurice, Laird & Ware, 2004) with an exchangeable correlation model to account for within hospital  $h_i$  correlation. The normal distribution  $N(\mathbf{0}, \sigma^2 \mathbf{I})$  represents a non-informative prior distribution (for small values of the precision  $\tau = 1/\sigma^2$ ) added to regularize cases with separable data.

## Analysis Module Output

In addition to the quantities  $\hat{\boldsymbol{\beta}}_Y, \hat{\boldsymbol{\beta}}_P, \hat{\pi}_X$  discussed above, the Analysis Module also calculates, for comparison purposes, the regression coefficients of the binomial model  $[Y | \mathbf{Z}]$  fitted using all the data, the binomial model  $[Y | \mathbf{X}]$  fitted using all the non-missing data, and the binomial model  $[Y | \mathbf{X}, P = 0]$  fitted using all the non-missing data with  $P = 0$ .



## The Prediction Module

The purpose of the Prediction Module (PM) is to predict, for each discharge record, the expected value of the adverse health outcome. These predictions are based on: i) the user's input dataset containing the same information, and having the same format as the analysis input dataset; and ii) a set of regression coefficients previously fitted by the Analysis Module using the data from a reference population. Since the adverse health outcome is binary (either it is present or it is not), the expected value for each discharge can be viewed as the probability that the adverse health outcome would have occurred for that discharge. These calculations are straightforward when there are no missing data, but they require high dimensional sums when data are missing.

### Overview

If POA data are being ignored, then the relevant output from the PM is the expected value of  $Y|Z$ . This is calculated with a simple scalar product of regression coefficients and covariates. The calculation is performed in the PM, but it results in the same number that would be obtained using SAS PROC SCORE. There is no MCMC involved in its calculation.

IF POA data are being accounted for in the calculations, then the relevant output from the PM is the expected value of  $Y|X, P=0$ .

- a) For discharge records where POA is observed, this, too, is calculated with a simple scalar product of regression coefficients and covariates. The MCMC is not involved.
- b) For discharge records where POA is missing, the expected value is calculated using a Gibbs Sampler MCMC as described below.

### Outline of the MCMC algorithm to predict $[Y|X, P = 0]$ using records with missing data

Before the MCMC begins:

- Read the 2 by 2 binary tables  $[X_k|Z_k]$  and the estimated regression coefficients of the model  $[P|X]$  fitted before the MCMC analysis discussed in the previous section;
- Read the estimated regression coefficients  $\hat{\beta}$  of the model  $[Y, X, P = 0]$  fitted by the MCMC analysis discussed in the previous section.

MCMC loop:

1. Build joint distribution  $[X, P, Z] = [P|X][X|Z]$ ;
2. Use full conditional distribution  $[X|P, Z] = [X, P, Z]/[P, Z]$  to draw missing  $X$ s (Gibbs Sampling); write the drawn missing  $X$  values to the chain. These values can be referred to as *imputed data*.
3. Use full conditional distribution  $[P|X, Z] = [X, P, Z]/[X, Z]$  to draw missing  $P$ s (Gibbs Sampling); write the drawn missing  $P$  values to the chain. These values can be referred to as *imputed data*.

4. Calculate predicted probability of an adverse outcome using the fitted regression coefficient  $\hat{\beta}_{MCMC}$ , the available  $X$  and  $P$  data, where measured, and the last imputed data, where  $X$  and  $P$  are missing (see 2. and 3.) The predicted probability, when  $P = 0$ , is calculated according to a logistic regression as the inverse logit of the scalar product  $\hat{\beta}_{MCMC} \cdot X$ .
5. Write the predicted probabilities of an adverse outcome when  $P = 0$  to the MCMC chain.
6. Go to 1 until total iterations equals that specified in the input XML file.

Note: points 1 through 3 of this section are similar to points 1 through 3 of the Analysis Module MCMC, only here we do not have  $Y$  data, which is what we are predicting.

During the MCMC loop:

- Drop burn-in entries
- Thin the chain, if appropriate (e.g., if the MCMC missing  $X$ s, missing  $P$ s, and the regression coefficients  $\beta$  are correlated in the chain)

After the MCMC loop:

- Calculate the expected values (average) of the components of the MCMC chain representing the predicted adverse outcome when  $P = 0$ ;

Note 1. The MCMC expected values are an unbiased estimated of the predicted adverse outcome assuming that no value is missing;

Note 2. The random numerical relative error introduced by the finiteness of the MCMC chain is inversely proportional to the square root of number of MCMC steps, and it becomes negligible compared to the statistical error of the predictions as the number of MCMC increases. See the 2011 report on Prediction Model accuracy posted on the AHRQ website.

<http://www.qualityindicators.ahrq.gov/Modules/Default.aspx>

## More Detailed Statistical Model

Let  $\hat{\beta}_Y, \hat{\beta}_P, \hat{\pi}_X$  be the regression coefficients fit by the AM as described in the previous section, and set

$$p(X', P') = [X', P' | Y, X', P', Z; \hat{\beta}_Y, \hat{\beta}_P, \hat{\pi}_X] \propto \\ \propto [Y | X', P'; \hat{\beta}_Y] \times [P' | X'; \hat{\beta}_P] \times [X' | Z; \hat{\pi}_X] \times [X' | X] \times [P' | P]$$

The main goal of the Prediction Module is to calculate

$$\Pr[Y_i = 1 | X_i, P_i = 0]$$

where we explicitly use the index  $i$  to indicate that the prediction is performed at the discharge record. For a record where both  $P_i$  and  $X_i$  are measured and  $P_i = 0$ , the predicted probability is simply given by

$$\Pr[Y_i = 1 | \mathbf{X}_i, P_i = 0] = \hat{\pi}_Y(\mathbf{X}_i) \equiv \text{logit}^{-1}(\mathbf{X}_i \hat{\boldsymbol{\beta}}_Y)$$

If  $P_i$  is missing, then we calculate the expected value of  $\hat{\pi}_Y(\mathbf{X}_i)(1 - P_i')$  over the distribution of the missing data  $p(\mathbf{X}'_i, P'_i)$ , namely

$$E_i[\hat{\pi}_Y] = \sum_{P'_i \in \{0,1\}} \hat{\pi}_Y(\mathbf{X}_i)(1 - P'_i) p(\mathbf{X}'_i, P'_i) = \hat{\pi}_Y(\mathbf{X}_i) p(\mathbf{X}_i, 0) = \text{logit}^{-1}(\mathbf{X}_i \hat{\boldsymbol{\beta}}_Y) p(\mathbf{X}_i, 0)$$

which is quick to compute. The general case however, where  $P_i$  and/or any combination of components of the vector  $\mathbf{X}_i$  is missing, requires the sum over all the possible combinations of missing values:

$$\begin{aligned} E_i[\hat{\pi}_Y] &= \sum_{P'_i, \mathbf{X}'_i} \hat{\pi}_Y(\mathbf{X}'_i)(1 - P'_i) p(\mathbf{X}'_i, P'_i) = \\ &= \sum_{P'_i, \mathbf{X}'_i} \text{logit}^{-1}(\mathbf{X}'_i \hat{\boldsymbol{\beta}}_Y)(1 - P'_i) p(\mathbf{X}'_i, P'_i) \end{aligned} \quad (\text{C.9})$$

Following the same argument used in the previous section, as the number of components of the vector of covariate  $\mathbf{X}$  increases, the deterministic sum quickly becomes unfeasible and an alternative approach is necessary. In this case, we evaluated the multidimensional sum using a Gibbs sampling implementation of the Importance Sampling Monte Carlo integration method (see chapter 7, paragraphs 7.6, 7.7 of the celebrated Numerical Recipes book (Press et al., 1992) for a primer introduction on Monte Carlo integration, references (Hammersley & Handscomb, 1964; Ripley, 1987; Rubinstein, 1981) for a deeper discussion, or many of the papers on the subject that can be freely found online.)

The methods works as follows: we draw a sample of imputed  $\mathbf{X}'_i, P'_i$  values from the distribution  $p(\mathbf{X}'_i, P'_i)$ , namely

$$(\mathbf{X}'_{i,s}, P'_{i,s}) \sim p(\mathbf{X}'_i, P'_i) \quad s = 1, \dots, N$$

using Gibbs Sampling to sample  $\mathbf{X}'_i$  and  $P'_i$  discussed in the Analysis Module section, then we approximate the sum (C.9) with the sample sum

$$I_N = \frac{1}{N} \sum_{s=1}^N \hat{\pi}_Y(\mathbf{X}'_{i,s})(1 - P'_{i,s})$$

Because Gibbs sampling generates a Markov chain, this method can be considered a MCMC method.

The numerical approximation of the Monte Carlo integration is known to be controlled by the sample variance

$$V_N = \frac{1}{N-1} \sum_{s=1}^N \left( \hat{\pi}_Y(\mathbf{X}'_{i,s})(1 - P'_{i,s}) \right)^2 - \frac{N}{N-1} \left( \frac{1}{N} \sum_{s=1}^N \hat{\pi}_Y(\mathbf{X}'_{i,s})(1 - P'_{i,s}) \right)^2$$

Since the distribution  $p$  has compact support and the function  $\hat{\pi}_Y(\mathbf{X}_i)$  is bounded, then the variance  $V_N$  is also bounded. Therefore, under the assumption that the sample  $(\mathbf{X}'_{i,s}, P'_{i,s})$  is ergodic (i.e. random), it follows from the central limit theorem that

$$I_N \rightarrow E_i[\hat{\pi}_Y]$$

in a probabilistic sense with a standard error equal to

$$\sigma_N = \sqrt{V_N/N}$$

The value  $V_N$  can be calculated together with  $I_N$  to provide an estimate of the Monte Carlo approximation error. However, regardless of  $V_N$ , the error of the MCMC integration scales as  $1/\sqrt{N}$ .

The PM also calculates, for comparative purposes, the expected values of the predictor  $\hat{\pi}_Y$  for the different sets of coefficients  $\hat{\beta}_Y$  estimated in the Analysis Module, the expected values of the predictor  $\hat{\pi}_P$ , and the marginal posterior probability of  $P'_i = 1$  given by

$$\sum_{\mathbf{X}'_i} p(\mathbf{X}'_i, 1)$$

## Appendix D. Helpful Background Information

This appendix includes some helpful information on both annual coding updates and software that is related to, or used by the AHRQ QI software. 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-9-CM and MS-DRG codes and revisions to existing codes. These changes are effective on October 1st. For example, Version 29 (fiscal year 2012) codes were effective October 1, 2011 and were incorporated in the version 4.4 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-9-CM procedure codes affect the Centers for Medicare and Medicaid Services (CMS) classification of “major operating room procedure” for postoperative PSIs and PDIs. Another use of ICD-9-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’s (CDC) National Center for Health Statistics (NCHS). Procedure and MS-DRG codes are maintained by the CMS. The activities of both agencies are conducted jointly through the ICD-9-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). It is anticipated that October 1, 2012 will witness only limited coding updates (from the September 14-15, 2011 and March 5, 2012 meetings of the Committee) to both the ICD-9-CM and ICD-10-CM/PCS codes to capture new technologies and diseases. The Committee meeting agendas and ICD-9-CM timeline is located at the [CMS site](#).

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/ICD10/11b14\\_2012\\_ICD10CM\\_and\\_GEMs.asp](http://www.cms.gov/ICD10/11b14_2012_ICD10CM_and_GEMs.asp) and [http://www.cms.gov/ICD10/11b15\\_2012\\_ICD10PCS.asp#TopOfPage](http://www.cms.gov/ICD10/11b15_2012_ICD10PCS.asp#TopOfPage)) web sites.

APR-DRG codes are maintained by 3M.

#### A.1 ICD-9-CM coding updates and coding guidelines

Information on ICD-9-CM coding updates is located on both the NCHS and CMS web sites: (<http://www.cdc.gov/nchs/icd/icd9cm.htm>)

([www.cdc.gov/nchs/data/icd/icd9cm\\_guidelines\\_2011.pdf](http://www.cdc.gov/nchs/data/icd/icd9cm_guidelines_2011.pdf))  
 ([http://www.cms.gov/ICD9ProviderDiagnosticCodes/01\\_overview.asp](http://www.cms.gov/ICD9ProviderDiagnosticCodes/01_overview.asp))

The anticipated coding updates for the subsequent version of the AHRQ QIs will consist of:

- New codes, if released.
- Limited ICD-9-CM coding revisions or deletions.
- NQF related updates, which may affect one or more indicators (This activity is performed in collaboration with task C.08. A set of NQF requested refinements have been submitted by AHRQ).

Activities during the base year will focus on these coding updates for the subsequent version of the AHRQ QIs. In general, updates to diagnosis and procedure codes are available on the NCHS or CMS web site. Preliminary updates are posted in March and final updates are posted in July. Diagnosis code updates are reported in Volume 1 (a tabular listing containing a numerical list of the disease code numbers) and Volume 2 (an alphabetical index to the disease entries). Procedure code updates are reported in Volume 3 (an alphabetic index and tabular list for surgical, diagnostic, and therapeutic procedures in hospitals and inpatient settings).

The meeting calendar of the Committee will be monitored on an ongoing basis for meeting status and updates to the meeting minutes, and the published coding changes (Volumes 1 and 2 for the diagnosis codes and Volume 3 for the procedure codes) and errata, both preliminary and final, will be reviewed.

The processes for evaluating the updates are described within each subsection below.

### **Diagnosis Codes**

An update consists of three documents.

- ICD-9-CM Index to Diseases Addenda – lists changes to the indexing of codes to diseases.
- ICD-9-CM Diagnosis Tabular Addenda – lists changes to the codes and code categories (defined as the first three digits).
- Conversion Table of New ICD-9-CM Codes – maps current codes to previous codes.

The update process consists of reviewing these documents to identify any coding changes that impact the numerator, denominator or exclusion logic of the AHRQ QI. There are two types of changes:

- A current code is split into two or more sub-codes and the current code is retired. Cases previously assigned to the current code are now assigned to the sub-codes.
- A new code or code category is created. Some cases previously assigned to a current code are now assigned to the new code.

Each change is evaluated to determine whether cases assigned to the codes belong in the numerator, denominator or exclusion logic of one or more AHRQ QI.

## **Procedure Codes**

An update consists of two documents.

- ICD-9-CM Procedure Tabular Addenda – lists changes to the codes and code categories
- Conversion Table of New ICD-9-CM Codes – maps current codes to previous codes.

The update process consists of reviewing these documents to identify any coding changes that impact the numerator, denominator or exclusion logic of the AHRQ QI. There are two types of changes.

- A current code is split into two or more sub-codes and the current code is retired. Cases previously assigned to the current code are now assigned to the sub-codes.
- A new code or code category is created. Some cases previously assigned to a current code are now assigned to the new code.

Each change is evaluated to determine whether cases assigned to the codes belong in the numerator, denominator or exclusion logic of one or more AHRQ QI.

## **A.2 DRG coding updates**

There are two editions of the DRGs. The first edition uses CMS-DRGs and the second edition uses MS-DRGs. The first edition is Version 24 and earlier; the second edition is Version 25 and later.

Updates to CMS-DRG are no longer supported by CMS.

Updates to MS-DRG codes are available on the CMS web site and in the Federal Register. Preliminary updates are posted in May and final updates or corrections are posted by August. (See <http://www.cms.gov/AcuteInpatientPPS>).

The update process consists of reviewing Table 5, which is a list of MS-DRGs, Relative Weighting Factors and Geometric and Arithmetic Mean Length of Stay and is one of the data tables from the fiscal year Inpatient Prospective Payment System from CMS. Ambiguity around the content of any update may usually be resolved through a review of the Federal Register notice. Prior to the implementation of the MS-DRGs, CMS would add and revise many DRGs annually. However, with the implementation of the MS-DRGs, changes are less frequent.

Activities during the base year will focus on reviewing the MS-DRG updates and determining what measure(s) are impacted with regards to the AHRQ QIs.

## **3M APR-DRG coding updates**

There is no public posting of updates to the APR-DRG. The commercial product is released in October with an update in April. A research license for the commercial product is available from AHRQ. The limited license grouper used in the AHRQ QI software is available on an ad hoc

basis under a voluntary arrangement with 3M. Contact information for the APR-DRG is as follows:

Anne M. Boucher  
 Implementation Manager  
 Clinical and Economic Research  
 3M Health Information Systems  
 100 Barnes Road  
 Wallingford, CT 06492  
 Telephone: (203) 949 6497  
 Email: [amboucher@mmm.com](mailto:amboucher@mmm.com)

Along with the limited license grouper, 3M provides documentation on changes to the APR-DRG logic. APR-DRG uses the same version numbering system used by NCHS and CMS. Prior to Version 23 (fiscal year 2006), 3M released a new version of the APR-DRG only once every five fiscal years with an ICD-9-CM mapping to maintain compatibility. Currently 3M releases a new version each fiscal year.

Updating the APR-DRG consists of the following steps:

1. Running the commercial product on the most recent year of Healthcare Cost and Utilization Project (HCUP) data available.
2. AHRQ has “pre-grouped” the HCUP data for selected states and made APR-DRG and risk-of-mortality subclass data elements available on the HCUP intramural databases. Step number 1 does not need to be done for these states.
3. Tabulating the frequency of APR-DRGs in the denominator of each IQI that uses the APR-DRG for risk-adjustment.
4. Retaining those APR-DRGs with at least 30 cases in the numerator.
5. Ensuring that those retained APR-DRGs are included in the covariate tables.

## **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 are also updated annually to reflect coding changes. The AHRQ QI support team does not independently review these changes; rather the coding changes are implemented without further review.

### **B.1 Comorbidity software**

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>).



The comorbidity software consists of two SAS programs. The first program, Creation of Format Library for Comorbidity Groups (Comformat.txt), creates a SAS format library that maps diagnosis codes into comorbidity indicators. Additional formats are also created to exclude conditions that may be complications or that may be related to the principal diagnosis. The second SAS program, Creation of Comorbidity Variables (Comoanaly.txt), applies the formats created above to a data set containing administrative data and then creates the comorbidity variables used to define the risk-adjustment covariates.

Updating the comorbidity software as used in the AHRQ QI software consists of the following steps:

- Comparing the current format program with the previous format program to identify any changes.
- Comparing the current analysis program with the analysis format program to identify any changes.
- Determine whether any of the changes present a problem for backwards compatibility and, if there is such a problem, design a solution.
- Implement any changes and solutions in the AHRQ QI software.

## **B.2 Clinical Classification Software (CCS)**

The CCS for ICD-9-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>).

Updating the clinical classification software as used in the AHRQ QI software consists of the following steps:

- Comparing the current format program with the previous format program to identify any changes.
- Determine whether any of the changes present a problem for backwards compatibility and, if there is such a problem, design a solution.
- Implement any changes and solutions in the AHRQ QI software.

## **B.3 Procedure classes**

The procedure classes assign ICD-9-CM procedure codes to one of four categories:

- Minor Diagnostic - Non-operating room procedures that are diagnostic.
- Minor Therapeutic - Non-operating 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>).

There is one file per fiscal year (PC.csv) that includes three elements: ICD-9-CM procedure codes, ICD-9-CM code labels, and procedure class assignments. In general, most of the changes relate to new procedure codes. However, sometimes the procedure class changes for an existing code. In these circumstances, the most recent assignment is used.

Updating the procedure classes as used in the AHRQ QI software consists of the following steps:

- Comparing the current procedure class assignments with the previous procedure class assignments to identify any changes.
- Special attention is given to operating room procedures in classes 3 and 4 (used to identify surgical discharges).
- Implement any changes in the AHRQ QI software.

## **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 maintain a consistent mapping between CMS DRGs and MS-DRGs, and 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 four digit code. The first two digits are the Major Diagnosis Category (MDC), and the second two digits are a sequence number (e.g., 01-04) within the MDC.

Updating the modified DRGs consists of the following steps:

- Identify the relevant AHRQ QIs for which the fiscal year MS-DRG changes apply. The MS-DRG changes are identified in the CMS Table 5 (a list of MS-DRGs, Relative Weighting Factors and Geometric and Arithmetic Mean Length of Stay) from the fiscal year Inpatient Prospective Payment System.
- Use the CMS crosswalk to pool CMS-DRGs and MS-DRGs into a single MDRG and compare with the MDRG categories table in the relevant risk adjustment tables document.
- Implement any changes in the AHRQ QI software.

### **C.2 Birth weight (BWHTCAT)**

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

Updating the birth weight categories consists of the following steps:

- Identify the relevant ICD-9-CM coding updates that pertain to the definition of the birth weight categories.
- Update the specifications, appendix and change log for the PDIs.
- Implement any changes in the AHRQ QI software.

### **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-9-CM codes (Original source Phibbs, et. al.<sup>5</sup>). Occasionally new codes are derived from existing codes.

Updating the CONGCATs consists of the following steps:

- Identify the relevant ICD-9-CM coding updates that pertain to the definition of the congenital anomalies.
- Update the specifications and change log for the relevant AHRQ QIs.
- Implement any changes in the AHRQ QI software.

### **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:

- Identify the relevant ICD-9-CM coding updates that pertain to the definition of the classifications.
- Update the specifications, appendix and change log for the relevant AHRQ QIs.
- 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. The RACHS-1 software is maintained on an ad hoc basis by Children's Hospital in Boston.

---

<sup>5</sup> Phibbs CS, Baker LC, Caughey AB, Danielsen B, Schmitt SK, Phibbs RH. Level and volume of neonatal intensive care and mortality in very-low-birth-weight infants. *New England Journal of Medicine*. 2007;356(21):2165-2175 & Supplement.

(See <http://www.ncbi.nlm.nih.gov/pubmed/15283367>).

Updating the RACHS-1 software consists of confirming the coding updates that apply to RACHS-1 from the Children's Hospital in Boston. The RACHS-1 stratifications should be added to the risk adjustment documentation under C.12.4.6 in the C.12 work plan.



# **2013 POPULATION FILE FOR USE WITH AHRQ QUALITY INDICATORS™ Version 4.5**

**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. HHSA290201200001C**

**Prepared by:**

**Battelle**  
505 King Avenue  
Columbus, OH 43201

**May 2013**

## Table of Contents

1.0 Overview.....	1
2.0 Data and Methodology.....	2
2.1 Census Data Files.....	2
2.1.1 Notable Differences of Population Estimates from 2000 Census to 2010 Census.....	4
2.1.2 Modifications to Census Estimates for use in the POP95T13.txt File.....	4
2.1.3 Census Data File Mapping to AHRQ QI Population File.....	5
2.2 Version History.....	5
2.2.1 Comparison of v4.4 and v4.5.....	6
3.0 POP95T13.txt File Specification.....	8

### List of Tables

Table 1. AHRQ QI Area-Level Indicators.....	1
Table 2. Census Dataset Descriptions and Sources.....	3
Table 3. Race Category Aggregations Based on Census Reporting Categories.....	4
Table 4. Population Files Used With Various Versions of AHRQ QI Software.....	6
Table 5. Data Fields in POP95T13.txt.....	10

### List of Figures

Figure 1. Comparison of population estimates for v4.4 and v4.5 by gender, age group and race for Madison County, Iowa (left) and Morehouse Parish, Louisiana (right) for the year 2009. The blue dashed lines represent $\pm 10\%$ deviation from the one-to-one line (red line).....	7
Figure 2. Comparison of population estimates for v4.4 and v4.5 by gender, age group and race for Miami-Dade County, Florida (left) and Maricopa County, Arizona (right) for the year 2009. The blue dashed lines represent $\pm 10\%$ deviation from the one-to-one line (red line). ....	8

## 1.0 Overview

The Agency for Healthcare Research and Quality (AHRQ) Quality Indicators (QI)™ include 36 area-level indicators (Table 1). These indicators are intended to measure healthcare quality across the population in a geographic area rather than for a single facility or provider. With a few exceptions, as noted in Table 1, the denominators for area-level indicators are the population of the area being examined, subset by age or (for some indicators) gender. The denominators for these indicators must be constructed from an outside source rather than drawn from a subset of discharges in the user's input file.

The objective of this document is to describe how the population data estimates are derived from public-use census data for use with the SAS QI Software Version 4.5 (SAS QI v4.5) and Windows QI Software Version 4.5 (WinQI v4.5). Population figures through 2013 for use with SAS QI v4.5 are provided in the file POP95T13.txt, available as a separate download on the AHRQ QI website. Population data are built into the installation package for WinQI v4.5.

**Table 1. AHRQ QI Area-Level Indicators**

IQI #26 Coronary Artery Bypass Graft (CABG) Rate	PQI #10 Dehydration Admission Rate
IQI #27 Percutaneous Coronary Intervention (PCI) Rate	PQI #11 Bacterial Pneumonia Admission Rate
IQI #28 Hysterectomy Rate	PQI #12 Urinary Tract Infection Admission Rate
IQI #29 Laminectomy or Spinal Fusion Rate	PQI #13 Angina Without Procedure Admission Rate
PSI #21 Retained Surgical Item or Unretrieved Device Fragment Rate	PQI #14 Uncontrolled Diabetes Admission Rate
PSI #22 Iatrogenic Pneumothorax Rate	PQI #15 Asthma in Younger Adults Admission Rate
PSI #23 Central Venous Catheter-Related Blood Stream Infection Rate	PQI #16 Lower-Extremity Amputation among Patients with Diabetes Rate
PSI #24 Postoperative Wound Dehiscence Rate	PQI #90 Prevention Quality Overall Composite
PSI #25 Accidental Puncture or Laceration Rate	PQI #91 Prevention Quality Acute Composite
PSI #26 Transfusion Reaction Rate	PQI #92 Prevention Quality Chronic Composite
PSI #27 Perioperative Hemorrhage or Hematoma Rate	PDI #14 Asthma Admission Rate
PQI #1 Diabetes Short-Term Complications Admission Rate	PDI #15 Diabetes Short-Term Complications Admission Rate
PQI #2 Perforated Appendix Admission Rate <sup>1</sup>	PDI #16 Gastroenteritis Admission Rate
PQI #3 Diabetes Long-Term Complications Admission Rate	PDI #17 Perforated Appendix Admission Rate <sup>1</sup>
PQI #5 Chronic Obstructive Pulmonary Disease (COPD) or Asthma in Older Adults Admission Rate	PDI #18 Urinary Tract Infection Admission Rate
PQI #7 Hypertension Admission Rate	PDI #90 Pediatric Quality Overall Composite
PQI #8 Heart Failure Admission Rate	PDI #91 Pediatric Quality Acute Composite
PQI #9 Low Birth Weight Rate <sup>1</sup>	PDI #92 Pediatric Quality Chronic Composite

<sup>1</sup>These indicators use discharge data from the input data file to estimate the denominator rather than demographic data from the population file.

## 2.0 Data and Methodology

Every year, the Census Bureau releases postcensal population estimates<sup>†</sup> (as of July 1 of each year) that are generated with the assistance of the Federal State Cooperative Program for Population Estimates (FSCPE) using residence, total births, total deaths, and net migration. With each new issue of July 1 estimates from the Census Bureau, the Census Bureau makes revisions to all years back to the last decennial census. Each decade, after a decennial census, the Census Bureau produces a set of intercensal estimates that provide annual population estimates that are adjusted to smooth the transition from one decennial census to the next. These estimates are used to derive the AHRQ QI Population File to be used with the AHRQ QI software.

### 2.1 Census Data Files

Public-use files of intercensal and postcensal estimates of county-level population by five-year age group, sex, race, and Hispanic origin were acquired from the Census Bureau (<http://www.census.gov/popest/>) covering the years 1995 through 2011. Table 2 presents detailed information and sources for the specific files acquired and used to generate the POP95T13.txt file for use within the AHRQ QI software.

---

<sup>†</sup> “Estimates are for the past, while projections are based on assumptions about future demographic trends. Estimates generally use existing data collected from various sources, while projections must assume what demographic trends will be in the future” (<http://www.census.gov/population/www/projections/aboutproj.html>)



**Table 2. Census Dataset Descriptions and Sources.**

DATA NAME	YEARS	BASE DECENNIAL YEAR	TYPE	SOURCE
Intercensal Estimates of the Resident Population by Five-Year Age Groups, Sex, Race, and Hispanic Origin for Counties	2000-2010	2010	Intercensal	<a href="http://www.census.gov/popest/data/intercensal/county/CO-EST00INT-alldata.html">http://www.census.gov/popest/data/intercensal/county/CO-EST00INT-alldata.html</a>
Annual Estimates of the Resident Population by Age, Sex, Race, and Hispanic Origin for Counties	2010-2011	2010	Postcensal	<a href="http://www.census.gov/popest/data/counties/asrh/2011/CC-EST2011-alldata.html">http://www.census.gov/popest/data/counties/asrh/2011/CC-EST2011-alldata.html</a>
Intercensal Estimates of the Resident Population by Single Year of Age and Sex for States and the United States	2000-2010	2010	Intercensal	<a href="http://www.census.gov/popest/data/intercensal/state/state2010.html">http://www.census.gov/popest/data/intercensal/state/state2010.html</a>
State Single Year of Age and Sex Population Estimates	2010-2011	2010	Postcensal	<a href="http://www.census.gov/popest/data/state/asrh/2011/index.html">http://www.census.gov/popest/data/state/asrh/2011/index.html</a>
State and County Intercensal Estimates by Demographic Characteristics	1990-1999	2000	Intercensal	<a href="http://www.census.gov/popest/data/intercensal/st-co/characteristics.html">http://www.census.gov/popest/data/intercensal/st-co/characteristics.html</a>

### 2.1.1 Notable Differences of Population Estimates from 2000 Census to 2010 Census

There are four counties that existed for the 2000 Census, but not for the 2010 Census (<http://www.census.gov/geo/www/tiger/tgrshp2010/usernotes.html>):

- 02201 - Prince of Wales-Outer Ketchikan Census Area, AK
- 02232 - Skagway-Hoonah-Angoon Census Area, AK
- 02280 - Wrangell-Petersburg Census Area, AK
- 51560 - Clifton Forge city, VA

In the 2010 Census, the populations from these four counties are distributed to other surrounding counties. This means that while the POP95T13.txt file contains estimates for these four defunct counties for the years 1995-1999, the POP95T13.txt file estimates for the years 2000-2013 are listed as “0” since they are based on 2010 Census county boundaries.

### 2.1.2 Modifications to Census Estimates for use in the POP95T13.txt File

Modifications to the census estimates were required to fit the specifications of the AHRQ QI software. The first is the categorization of race and Hispanic origin. Table 3 depicts how the race categories used by the AHRQ QI software were defined from the census race and Hispanic origin groupings. This set of race categorizations captures the entire US population.

**Table 3. Race Category Aggregations Based on Census Reporting Categories.**

RACE CATEGORY	DESCRIPTION
1	Non-Hispanic, White Alone
2	Non-Hispanic, Black Alone
3	Hispanic
4	Non-Hispanic, Asian Alone OR Non-Hispanic, Native Hawaiian and Other Pacific Islander Alone
5	Non-Hispanic, American Indian and Alaska Native Alone
6	Non-Hispanic, Two or More Races

In addition, the population of interest for the area-level indicators in the Pediatric Quality Indicator (PDI) module is the population ages 17 and under, while the population of interest for the other indicator modules is the population ages 18 and older. The default five-year age groups reported by the Census Bureau are 15-19 years of age and 20-24 years of age. To capture the separation between the pediatric and adult populations, the POP95T13.txt file contains an age range that spans the ages of 18-24 that is constructed using the two default census age groups. To generate the 18-24 year old age group, state-level estimates of population by sex and single year of age (see Table 2) were used to calculate the percent of the population between 15 and 19 years old (the age grouping for the county-level data) that are between 18 and 19 years old. Then, the county-level population of 18-19 year olds was subtracted from the census-defined age group of 15-19 (to form the 15-17 age group) and added to the 20-24 age group (to form the 18-24 age group).

### 2.1.3 Census Data File Mapping to AHRQ QI Population File

The POP95T13.txt file population estimates for 1995 through 1999 are based on intercensal estimates by demographic characteristics (Table 2). Since these data are adjusted to the 2000 Census, they are no longer updated by the Census Bureau with more recent postcensal estimates and the estimates are unchanged from version release to version release of the AHRQ QI software.

The POP95T13.txt file population estimates for 2000 through 2010 are based on intercensal estimates by demographic characteristics that are adjusted to the 2010 Census. The POP95T13.txt file population estimates for 2011 are based on postcensal estimates by demographic characteristics that use the 2010 Census as the base.

Public-use files of postcensal population estimates from the Census Bureau are currently available only through 2011. The POP95T13.txt file contains population estimates for 2012 and 2013 based on linear projections of the population counts for each county, sex, age group, and race combination. The projections were made according to the following model:

$$y_{ijt} = \alpha_{ij} + \beta_{ij}t,$$

where  $i$  is the county (1, 2, ..., 3147),  $j$  is an indicator of demographics representing a combination of sex, age group, and race (1, 2, ..., 216), and  $t$  is the year (2000, 2001, ..., 2011). That is, we fit a county-specific linear growth model for each demographic group. The population estimates for each county and demographic combination,  $\hat{y}$ , for 2012 and 2013 were calculated using the following equations:

$$\hat{y}_{ij2012} = \hat{\alpha}_{ij} + \hat{\beta}_{ij}2012$$

$$\hat{y}_{ij2013} = \hat{\alpha}_{ij} + \hat{\beta}_{ij}2013.$$

where  $\hat{\alpha}_{ij}$  and  $\hat{\beta}_{ij}$  are the coefficients estimated from the linear regression models.

## 2.2 Version History

The population file released with each version of the software is generated with the most recent data available at the time of software development. As such, this file will change from version to version (including the filename) as data are updated and released by the Census Bureau. The differences between population files for AHRQ QI software release versions can be caused by changes in population estimates themselves and/or changes in methodology. Table 4 summarizes the population files for AHRQ QI software release versions. Note that data for population files included with previous releases of the AHRQ QI software are not updated with each new release.

**Table 4. Population Files Used With Various Versions of AHRQ QI Software**

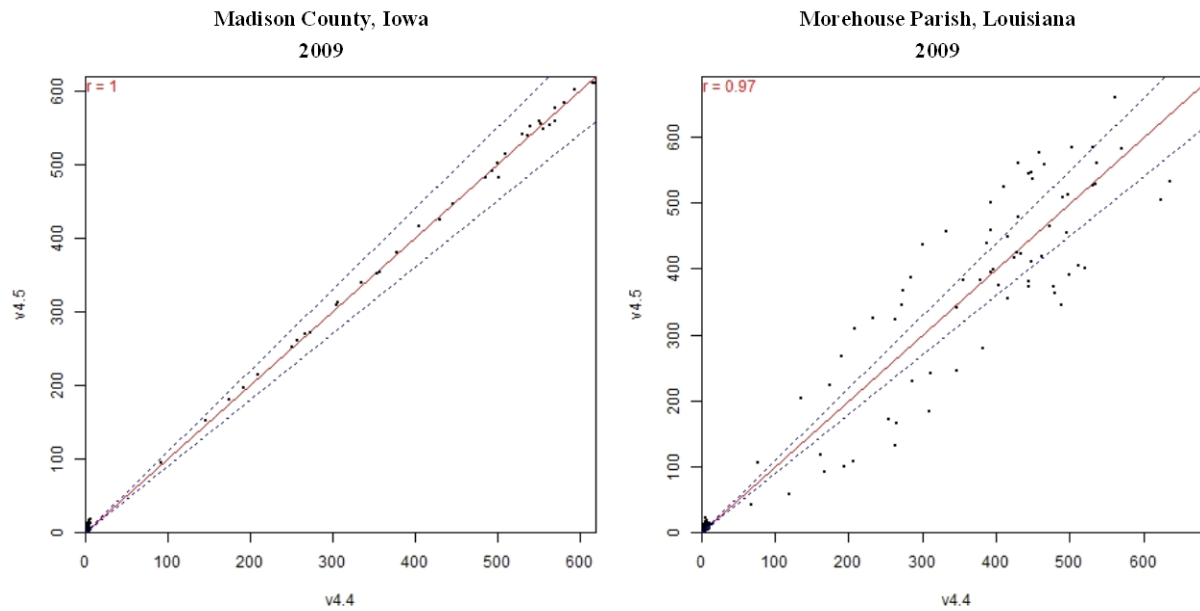
SOFTWARE RELEASE (FILENAME)	YEARS	BASE DECENNIAL YEAR	DATA SUMMARY	METHODOLOGY SUMMARY
v4.5 (POP95T13.TXT)	Estimates: 1995-1999	2000	(1) Sex/Age/Race by County (2) Age 18-24 by State	Permutated file of sex/age/race by county Used state estimate of population from 18-24 to break 15-19 and 20-24 age groups into 15-17 and 18-24
	Estimates: 2000-2011 Projections: 2012-2013	2010	(1) Age/Sex/Race by County (2) Age (single year) by State	Permutated file of sex/age/race by county Used state estimate of single year of age to break 15-19 and 20-24 age groups into 15-17 and 18-24
v4.4 (POP95T12.TXT)	Estimates: 1995-1999	2000	(1) Sex/Age/Race by County (2) Age 18-24 by State	Permutated file of sex/age/race by county Used state estimate of population from 18-24 to break 15-19 and 20-24 age groups into 15-17 and 18-24
	Estimates: 2000-2010 Projections: 2011-2012	2010	(1) Sex/Age by County (2) Sex/Race by County (3) Age (single year) by State	Combined sex/age and sex/race files by county to get estimates of sex/age/race Used state estimate of single year of age to break 15-19 and 20-24 age groups into 15-17 and 18-24
v4.3 (POP95T11.TXT)	Estimates: 1995-2009 Projections: 2010-2011	2000	(1) Sex/Age/Race by County (2) Age 18-24 by State	Permutated file of sex/age/race by county Used state estimate of population from 18-24 to break 15-19 and 20-24 age groups into 15-17 and 18-24

### 2.2.1 Comparison of v4.4 and v4.5

At the time of the AHRQ QI v4.4 development, the Census Bureau had not yet released the intercensal estimates of population by age, sex, race, and Hispanic origin at the county level for the years 2000 through 2010 that were updated to be consistent with the 2010 Census. In order to use the most recent data available, two separate, county-level files (one containing sex and age and the other containing sex and race) were merged to generate the estimates by sex, age, and race. In this fashion, the distribution of age group categories was applied evenly across all race categories (e.g. the percentage of non-Hispanic white males estimated to be between 0-4 years old was equal to the percentage of Hispanic males estimated to be between 0-4 years old for a given county). A comparison of the v4.4 and v4.5 files (with the v4.5 files based on updated 2010 Census data) revealed that this assumption is not necessarily true for all counties and races.

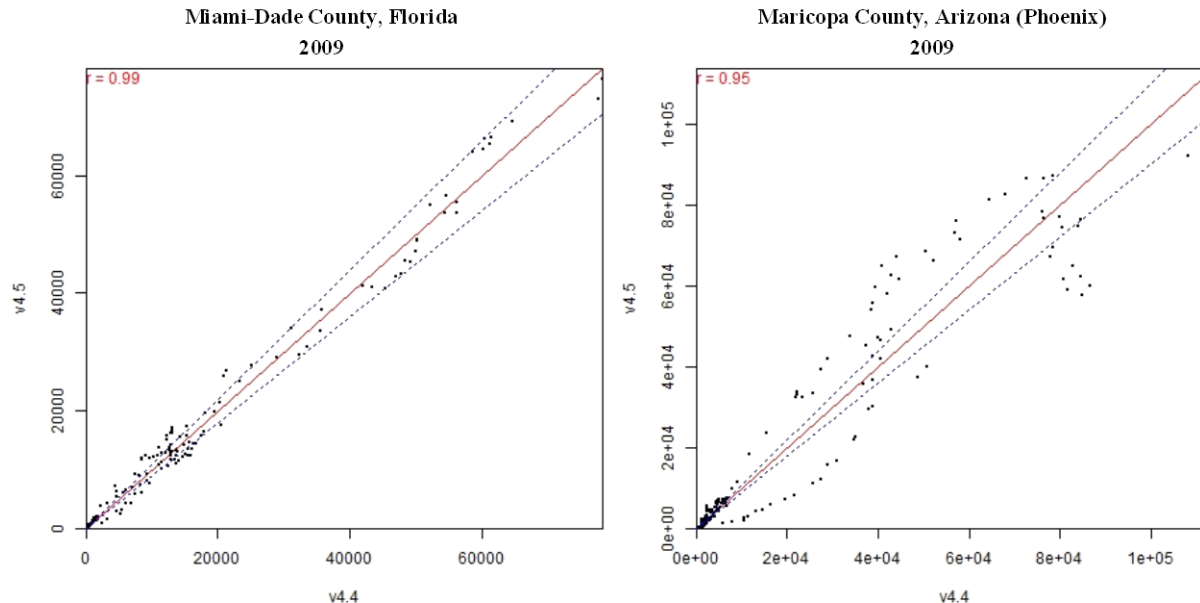
For counties where the age group by race distribution is approximately equal to the total age group distribution (i.e., not race dependent), there are not large differences between the population file used in v4.4 and that used in v4.5. However, for counties that have significantly different age group distributions for different races, large differences may be observed. For example, Figure 1 shows a comparison of the population estimates for two small counties (total populations less than 30,000). Each individual symbol (n=216) on the plot represents a gender, age group, race observation for the county. The blue, dashed lines indicate a  $\pm 10\%$  deviation from the one-to-one line indicating perfect agreement between the v4.4 and v4.5 estimates.

Madison County, Iowa has a predominantly non-Hispanic white population (>97% in 2009), resulting in estimates that agree very well between the two population files, while the estimates for Morehouse Parish, Louisiana, which has almost equal non-Hispanic white (51%) and African American (47%) populations, demonstrate some large differences. These differences between v4.4 and v4.5 occur because the two race categories in this county have different age group distributions, while the v4.4 methodology applied a single distribution across all races.



**Figure 1. Comparison of population estimates for v4.4 and v4.5 by gender, age group and race for Madison County, Iowa (left) and Morehouse Parish, Louisiana (right) for the year 2009. The blue dashed lines represent  $\pm 10\%$  deviation from the one-to-one line (red line).**

While the previous example was for two small counties, similar results are observed for large counties. Figure 2 shows the comparison of population estimates by gender, race, and age group for Miami-Dade County, Florida and Maricopa County, Arizona (Phoenix). The estimates for Miami-Dade County are more similar between population file versions than for Maricopa County. The age group distributions for the three races that contribute most to the total population of Miami-Dade County, Hispanic (65%), African American (17%) and non-Hispanic white (16%) all follow a very similar pattern, resulting in smaller deviations in the v4.5 estimate from the v4.4 estimate. However, in Maricopa County, the distributions for the two largest-contributing race groups, non-Hispanic whites (59%) and Hispanics (29%), follow very different patterns, meaning that the age group distribution applied in v4.4 for all races was not representative of the age group distributions for both of these races.



**Figure 2. Comparison of population estimates for v4.4 and v4.5 by gender, age group and race for Miami-Dade County, Florida (left) and Maricopa County, Arizona (right) for the year 2009. The blue dashed lines represent  $\pm 10\%$  deviation from the one-to-one line (red line).**

To summarize, the largest effects of this change are realized in large counties that have two or more race groups that contribute large proportions to the total population, but have different age group distributions, such as Maricopa County, Arizona and Los Angeles, California (not shown here). Small counties with similar demographics (e.g. Morehouse Parish, Louisiana) will also see a difference, though the absolute differences (i.e. numbers of people) are not as large. Counties that have one dominant race category and those with more than one that have very similar age group distributions will still have changes to the population estimates, but they are likely to be relatively small.

### 3.0 POP95T13.txt File Specification

The POP95T13.txt file is an ASCII-based text file containing 679,752 records with a fixed logical record length of 150 bytes. It is in fixed column format. Table 5 presents the file's specific fields and the code schema used for each field.

The file is structured for use with AHRQ QI programs PQSASA2.SAS, PQSASA3.SAS, PSSASA2.SAS, IQSASA2.SAS, IQSASA3.SAS, PDSASA2.SAS, and PDSASA3.SAS, as well as the Windows QI (WinQI) software. As such, any modification to this file will affect the operation of these programs.

A given county is identified by the Federal Information Processing Standards code (FIPS code) for the state in which it is located and by the county's FIPS code. For each county within the U.S., the file contains 216 records: a record for each unique combination of gender, eighteen age groups, and six race groups. Each physical record represents a gender, age group, and race

group combination for that county and contains population estimates (rounded to integer values) for that combination for each year from 1995 through 2013.

The file has data for 3,147 counties or “equivalent areas”, defined to constitute primary divisions of their states. “Equivalent areas” include the independent cities of Baltimore, Maryland; St. Louis, Missouri; Carson City, Nevada; and 39 independent cities in Virginia. Because they are independent of any contiguous county, they are treated as separate counties with their own population records. Population figures for surrounding counties exclude them. Differences in the record count from previous population files are due to changes in county definitions or such independent cities. Definitions for state and county FIPS codes can be found at [http://quickfacts.census.gov/qfd/meta/long\\_fips.htm](http://quickfacts.census.gov/qfd/meta/long_fips.htm).

**Table 5. Data Fields in POP95T13.txt**

FIELD	VARIABLE	COLUMN POSITION	FORMAT	CODES
1	State	1-2	Zero Filled Numeric	FIPS Code
2	County	3-5	Zero Filled Numeric	FIPS Code
3	Sex	7	Numeric	1=Male, 2=Female
4	Age Group	9-10	Numeric	1=0-4 years 2=5-9 years 3=10-14 years 4=15-17 years 5=18-24 years 6=25-29 years 7=30-34 years 8=35-39 years 9=40-44 years 10=45-49 years 11=50-54 years 12=55-59 years 13=60-64 years 14=65-69 years 15=70-74 years 16=75-79 years 17=80-84 years 18=85+ years
5	Race	12	Numeric	1=White, 2=Black, 3=Hispanic, 4=Asian & PI, 5=Amer. Indian, 6=Other
6	1995 Population	13-19	Numeric	Integer Totals
7	1996 Population	20-26	Numeric	
8	1997 Population	27-33	Numeric	
9	1998 Population	34-40	Numeric	
10	1999 Population	41-47	Numeric	
11	2000 Population	48-54	Numeric	
12	2001 Population	55-61	Numeric	
13	2002 Population	62-68	Numeric	
14	2003 Population	69-75	Numeric	
15	2004 Population	76-82	Numeric	
16	2005 Population	83-89	Numeric	
17	2006 Population	90-96	Numeric	
18	2007 Population	97-103	Numeric	
19	2008 Population	104-110	Numeric	
20	2009 Population	111-117	Numeric	
21	2010 Population	118-124	Numeric	
22	2011 Population	125-131	Numeric	
23	2012 Population	132-138	Numeric	
24	2013 Population	139-145	Numeric	



**AHRQ Quality Indicators**  
**Analytic Template**  
**Version 4.5**  
 (Last Updated 11/15/2013)

Measure #: PQI 14  
 Measure Name: Uncontrolled Diabetes Admission Rate

## I. Sample

The area universe is defined as the county of the residence of the patient for discharges in the hospital universe. The hospital universe is defined as all hospitals located in the U.S. that are open during any part of the calendar year and designated as community hospitals in the AHA Annual Survey Database (Health Forum, LLC © 2011). The AHA defines community hospitals as follows: "All non-Federal, short-term, general, and other specialty hospitals, excluding hospital units of institutions." Starting in 2005, the AHA included long term acute care facilities in the definition of community hospitals. These facilities provide acute care services to patients who need long term hospitalization (stays of more than 25 days). Consequently, Veterans Hospitals and other Federal facilities (Department of Defense and Indian Health Service) are excluded. Beginning in 1998, we excluded short-term rehabilitation hospitals from the universe because the type of care provided and the characteristics of the discharges from these facilities were markedly different from other short-term hospitals.

[http://hcup-us.ahrq.gov/db/nation/nis/NIS\\_Introduction\\_2011.pdf](http://hcup-us.ahrq.gov/db/nation/nis/NIS_Introduction_2011.pdf)

## II. Empirical Testing

### A. Reference Population

Table 1. Reference Population

Year/ Characteristic	Areas	Outcome of Interest	Population at Risk	Observed Rate Per 100,000
2011	3,112	41,449	236,853,218	17.500
2010	3,111	43,238	234,354,184	18.450
2009	3,112	44,327	231,837,796	19.120
2008	3,111	43,620	229,336,267	19.020
2007	3,107	41,024	226,777,940	18.090
<b>Performance Score Distribution 2011 (Rate per 100,000)</b>				
<b>5<sup>th</sup></b>	<b>25<sup>th</sup></b>	<b>Median</b>	<b>75<sup>th</sup></b>	<b>95<sup>th</sup></b>
3.282	8.604	14.801	23.491	40.940

Source: HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2007-2011. Agency for Healthcare Research and Quality, Rockville, MD. [www.hcup-us.ahrq.gov/sidoverview.jsp](http://www.hcup-us.ahrq.gov/sidoverview.jsp). (AHRQ QI Software Version 4.5)

## B. Reliability

Our metric of reliability is the signal to noise ratio, which is the ratio of the between area variance (signal) to the within area variance (noise). The formula is  $\text{signal} / (\text{signal} + \text{noise})$ . There is an area-specific signal to noise ratio, which is used as an Empirical Bayes univariate shrinkage estimator. The overall signal to noise ratio is a weighted average of the area-specific signal-to-noise ratio, where the weight is  $[1 / (\text{signal} + \text{noise})^2]$ . The signal is calculated using an iterative method. The analysis reports the reliability of the risk-adjusted rate (before applying the empirical Bayes univariate shrinkage estimator).

Table 2. Reliability by Area Size Decile

Size Decile	Number of Areas	Ave. Number of Persons per Area in Decile	Ave. Signal-to-Noise Ratio for Areas in Decile	Percent of Signal Variance Explained by Performance Score
1	312	2,278.3	0.38413	0.78676
2	311	5,657.7	0.61610	0.82285
3	311	8,817.0	0.71270	0.84669
4	311	12,640.8	0.77972	0.86806
5	311	17,288.9	0.82716	0.88670
6	312	23,989.2	0.86772	0.90574
7	311	33,767.9	0.90092	0.92394
8	311	53,199.9	0.93367	0.94494
9	311	103,760.9	0.96302	0.96689
10	311	500,100.9	0.98761	0.98817
Overall	3,112	76,109.6	0.96319	0.97232

Source: HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2011. Agency for Healthcare Research and Quality, Rockville, MD. [www.hcup-us.ahrq.gov/sidoverview.jsp](http://www.hcup-us.ahrq.gov/sidoverview.jsp). (AHRQ QI Software Version 4.5)

## C. Validity

We conduct construct validity testing to examine the association between the risk-adjusted rate and area structural characteristics potentially associated with quality of care, including prior performance, using regression analysis.

Table 3. Structure Measures Used to Estimate Prior Probability

Measure	How it is measured	Less Access to High Quality Outpatient Care Construct (F1)	Less Market Competition Construct (F2)
MD Density	Number of Physicians in Patient Care per Person	Areas with less physicians per person have less access to high quality outpatient care	Areas with more physicians per person have less market competition
Excess Capacity	Percent of Available Short-term General Hospital Beds per Total Beds	Areas with greater excess bed capacity have supply side incentive to have greater rates of admission	Areas with less excess bed capacity have less market competition
Poverty Status	Percent of Persons in Poverty	Areas with greater persons in poverty have less access to high quality outpatient care	Areas with greater persons in poverty have less market competition
Insurance Status	Percent of Persons (Under 65) without Health Insurance	Areas with greater persons without health insurance have less access to high quality outpatient care	Areas with greater persons without health insurance have less market competition
Population Density	Population Density per Square Mile	Areas with less population density have less access to high quality outpatient care	Areas with more population density have less market competition

Source: Area Health Resource File (AHRF) 2012-2013. US Department of Health and Human Services, Health Resources and Services Administration, Bureau of Health Professions, Rockville, MD.

Given the stated rationale, the expectation for the regression analysis given the expected relationship between the “Less Access to High Quality Outpatient Care” construct validity measure (F1) and the area risk-adjusted rate is a positive, statistically significant coefficient. The expectation for the regression analysis given the expected relationship between the “More Market Competition” construct validity measure (F2) and the area risk-adjusted rate is a positive, statistically significant coefficient

Table 4. Regression on Structure Measures

Variable	Label	Coef.	Std. Err	t	P> t	[95% Conf. Interval]
F1	Access to Quality Care	0.000064	0.000004	14.55	0.0000	0.00006
F2	Market Competition	0.000064	0.000007	9.17	0.0000	0.00005
_cons	Constant	0.000162	0.000003	50.20	0.0000	0.00016
F1	Access to Quality Care	0.000014	0.000003	4.50	0.0000	0.000008
F2	Market Competition	0.000011	0.000003	3.53	0.0000	0.000005
prior2	Prior Performance	0.728192	0.026613	27.36	0.0000	0.676011
_cons	Constant	0.000033	0.000004	7.51	0.0000	0.000025

Note: the dependent variable in the regression is the risk adjusted rate

#### D. Performance

We calculate the posterior probability distribution for each area parameterized using the Gamma distribution. We then calculate the probability that the area is better or worse than the reference population rate at a 95 percent probability overall and by area size decile. The analysis is with the computed performance scores for the measure as specified (including shrinkage estimator).

Table 5. Performance Categories by Area Size Decile

Size Decile	Number of Areas	Ave. Number of persons per Area in Decile	Proportion Better	Proportion Worse
1	312	2,278.3	0.10256	0.16026
2	311	5,657.7	0.21865	0.21222
3	311	8,817.0	0.31833	0.24116
4	311	12,640.8	0.34084	0.23794
5	311	17,288.9	0.35048	0.25723
6	312	23,989.2	0.35897	0.28526
7	311	33,767.9	0.48232	0.22508
8	311	53,199.9	0.51768	0.17363
9	311	103,760.9	0.56913	0.20257
10	311	500,100.9	0.60129	0.25402
	3,112	76,109.6	0.38593	0.22494
Patient weighted			0.51930	0.27117

Source: HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2011. Agency for Healthcare Research and Quality, Rockville, MD. [www.hcup-us.ahrq.gov/sidoverview.jsp](http://www.hcup-us.ahrq.gov/sidoverview.jsp). (AHRQ QI Software Version 4.5)

## E. Model Discrimination and Calibration

One calculates the c-statistic by taking all possible pairs of cases consisting of one case that experienced the event of interest and one case that did not experience the event of interest. The c-statistic is the proportion of such pairs in which the case that experienced the event had a higher predicted probability of experiencing the event than the case that did not experience the event.

Table 6. Model Discrimination and Calibration

<b>Predicted Rate Decile</b>	<b>Number of Persons per Decile</b>	<b>Predicted Rate</b>	<b>Observed Rate</b>
1	23,685,697	0.000050	0.000043
2	23,688,139	0.000065	0.000059
3	23,683,815	0.000089	0.000084
4	23,691,391	0.000126	0.000117
5	23,720,307	0.000173	0.000172
6	23,670,597	0.000227	0.000201
7	23,658,295	0.000243	0.000225
8	23,704,140	0.000259	0.000242
9	23,670,057	0.000288	0.000266
10	23,680,780	0.000363	0.000342
C-statistic	Not Calculated		

Source: HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2011. Agency for Healthcare Research and Quality, Rockville, MD. [www.hcup-us.ahrq.gov/sidoverview.jsp](http://www.hcup-us.ahrq.gov/sidoverview.jsp).

A model that is well calibration will have observed values similar to predicted values across the predicted value deciles. Although there are statistical tests of such “goodness of fit” the tests generally are not informative for datasets with large sample sizes.

## F. Forecasting

With respect to the persistence of the area risk adjusted rate, we conduct a descriptive analysis to examine the distribution of the current year risk-adjusted rate by the prior year performance score performance decile. The R-square is a statistic for the proportion of variation in the risk-adjusted rate captured by variation in the prior year performance score.

Table 7. Forecasting

<b>Prior Year Performance Score Quintile</b>	<b>Number of Areas Per Quintile</b>	<b>Prior Year Performance Score</b>	<b>Current Year Risk-adjusted Rate</b>
1	312	0.000018	0.000036
2	311	0.000048	0.000067
3	311	0.000075	0.000084
4	311	0.000100	0.000112
5	311	0.000129	0.000128
6	312	0.000158	0.000134
7	311	0.000194	0.000171
8	311	0.000243	0.000197
9	311	0.000323	0.000273
10	311	0.000609	0.000495
R-Squared		0.4263	

Source: HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2010-11. Agency for Healthcare Research and Quality, Rockville, MD. [www.hcup-us.ahrq.gov/sidoverview.jsp](http://www.hcup-us.ahrq.gov/sidoverview.jsp). (AHRQ QI Software Version 4.5)

## G. Preventability

Our metric of preventability is the proportion of events that are potentially preventable if patients selected areas performing at the level of the benchmark (i.e. the 20<sup>th</sup> percentile (better) in the probability distribution). The metric suggests that 62.0% of the events are potentially preventable.

Table 8. Preventability

Performance Score Decile	Ave. Performance Score	Number of Areas per Decile	Ave. Number of Persons per Area in Decile	Total Number of Persons in Decile	Total Events	Proportion Potentially Preventable Events	Potentially Preventable Events	Expected Value of Information
1	0.000033	311.2	76,109.6	23,685,321.8	791	0.000000	0.0	
2	0.000063	311.2	76,109.6	23,685,321.8	1,490	0.000000	0.0	
3	0.000087	311.2	76,109.6	23,685,321.8	2,066	0.000013	310.0	
4	0.000111	311.2	76,109.6	23,685,321.8	2,630	0.000037	874.5	
5	0.000136	311.2	76,109.6	23,685,321.8	3,224	0.000062	1,468.0	
6	0.000164	311.2	76,109.6	23,685,321.8	3,882	0.000090	2,126.2	
7	0.000197	311.2	76,109.6	23,685,321.8	4,655	0.000122	2,899.3	
8	0.000238	311.2	76,109.6	23,685,321.8	5,637	0.000164	3,880.8	
9	0.000299	311.2	76,109.6	23,685,321.8	7,074	0.000225	5,317.8	
10	0.000485	311.2	76,109.6	23,685,321.8	11,488	0.000411	9,731.8	
Overall		3,112	76,110	236,853,218	42,937	0.000112	26,608	
Proportion Preventable							0.6197	

Source: HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2011. Agency for Healthcare Research and Quality, Rockville, MD. (AHRQ QI Software Version 4.5)

## H. Information Value

The expected value of information (EVI) is the change in the expected number of potentially preventable events (“opportunity loss”) accounting for the uncertainty and low information context in the performance score. A negative EVI means that there is some uncertainty in the expected number of potentially preventable events, while a positive EVI means that the effective sample size might be increased. Ideally the expected value of information would be close to zero.

Table 9. Expected Value of Information

Performance Score Decile	Ave. Performance Score	Number of Areas per Decile	Ave. Number of Persons per Area in Decile	Total Number of Persons in Decile	Total Events	Proportion Potentially Preventable Events	Potentially Preventable Events	Expected Value of Information
1	0.000026	399	62,686.2	25,011,778	640	0.000002	50	-50
2	0.000062	368	74,217.1	27,311,893	1,696	0.000004	108	-108
3	0.000085	325	81,429.3	26,464,536	2,239	0.000014	365	-55
4	0.000110	313	73,667.9	23,058,062	2,536	0.000037	849	26
5	0.000136	305	80,864.2	24,663,584	3,362	0.000063	1,546	-78
6	0.000166	305	109,756.5	33,475,733	5,540	0.000092	3,073	-947
7	0.000197	303	66,061.0	20,016,471	3,943	0.000124	2,472	427
8	0.000239	252	87,856.9	22,139,949	5,287	0.000165	3,655	226
9	0.000305	215	83,561.0	17,965,615	5,485	0.000232	4,161	1,157
10	0.000473	327	51,209.8	16,745,601	7,922	0.000400	6,690	3,042
Overall		3,112	76,110	236,853,218	38,654	0.000097	22,975	3,633
Proportion Preventable							0.5944	0.1468

Source: HCUP State Inpatient Databases (SID). Healthcare Cost and Utilization Project (HCUP). 2011. Agency for Healthcare Research and Quality, Rockville, MD. (AHRQ QI Software Version 4.5)



Admissions for uncontrolled diabetes without complications<sup>a</sup> per 100,000 population, age 18 and over (PQI 14)

Adjusted rates by patient and hospital characteristics, 2011 (HCUPnet)

Patient/hospital characteristic	2011 Adjusted Rate <sup>b</sup>		P-value: Relative to marked group <sup>c</sup>
	Estimate	Standard error	
Total U.S.	19.811	0.697	
<b>Patient characteristic:</b>			
Age groups for conditions affecting any age			
18-44 <sup>c</sup>	9.920	0.443	
45-64	25.522	0.990	0.000
65 and over	35.834	1.393	0.000
Age groups for conditions affecting primarily elderly			
65-69 <sup>c</sup>	29.604	1.358	
70-74	37.072	1.929	0.002
75-79	40.262	2.162	0.000
80-84	40.976	2.402	0.000
85 and over	37.714	2.799	0.009
Gender:			
Male <sup>c</sup>	20.912	0.745	
Female	18.784	0.715	0.039
Median income of patient's ZIP Code:			
First quartile (lowest income)	35.076	2.018	0.000
Second quartile	20.303	0.992	0.000
Third quartile	15.127	0.864	0.000
Fourth quartile (highest income) <sup>c</sup>	9.879	0.778	
Location of patient residence (NCHS):			
Large central metropolitan	25.641	2.230	0.004
Large fringe metropolitan <sup>c</sup>	17.878	1.571	
Medium metropolitan	13.716	1.675	0.070
Small metropolitan	16.037	1.866	0.450
Micropolitan	19.559	1.765	0.477
Noncore	27.456	2.106	0.000
<b>Hospital characteristic:</b>			
Location of inpatient treatment:			
Northeast <sup>c</sup>	23.687	2.003	
Midwest	19.632	1.301	0.089
South	23.041	1.295	0.786
West	11.384	0.843	0.000

Source: Agency for Healthcare Research and Quality (AHRQ), Center for Delivery, Organization, and Markets, Healthcare Cost and Utilization Project, Nationwide Inpatient Sample, 2011, and AHRQ Quality Indicators, version 4.4.

<sup>a</sup> Consistent with the AHRQ PQI software, diabetes without complications must be the principal diagnosis and exclusions include transfers from other institutions.

<sup>b</sup> Rates are adjusted by age and gender using the total U.S. resident population for 2010 as the standard population; when reporting is by age, the adjustment is by gender only; when reporting is by gender, the adjustment is by age only.

<sup>c</sup> Reference for p-value test statistics.

NCHS - National Center for Health Statistics designation for urban-rural locations.