AHRQ

AHRQ QUALITY INDICATORS

Prevention Quality Indicators

SOFTWARE DOCUMENTATION

AHRQ Quality Indicators

Prevention Quality Indicators: Software Documentation, Version 3.0a SAS

Department of Health and Human Services
Agency for Healthcare Research and Quality
http://www.qualityindicators.ahrq.gov

October 2001
Version 3.0a (February 20, 2006)

Abstract

The value of information on health care quality has never been so widely recognized; yet many organizations lack the resources and/or expertise to build a quality information program from the ground up. Recognizing this, the AHRQ Quality Indicators were developed specifically to meet the short-term needs for information on health care quality using standardized, user-friendly methods and existing sources of data.

This module of the AHRQ Quality Indicators was designed to capitalize on the availability of administrative data on inpatient stays to produce information about 14 Prevention Quality Indicators (PQIs): *potentially avoidable hospitalizations for ambulatory care sensitive condition* (ACSC) indicators, which involve admissions that evidence suggest could have been avoided, at least in part, through better access to high-quality outpatient care.

This document is the software documentation for the Prevention Quality Indicators software Version 3.0a, which is provided on the AHRQ Web site.

The software was developed in SAS, for use on a personal computer. By making this tool available, we hope to assist others in producing information on health care quality more cost effectively.

Details on the development of the Prevention Quality Indicators can be found in *Guide to Prevention Quality Indicators: Hospital Admission for Ambulatory Care Sensitive Conditions* available at http://www.qualityindicators.ahrq.gov/pqi_download.htm.

Acknowledgments

Support efforts, including refinement and enhancement of the AHRQ Quality Indicators and related products, are provided by the Support for Quality Indicators-II contract team.

The following individuals from **Battelle Memorial Institute, Stanford University,** and **University of California (UC)** constitute the Support for Quality Indicators-II core team:

Sheryl M. Davies, M.A.
Bruce Ellis, M.S.
Jeffrey Geppert, J.D.
Mark Gritz, Ph.D.
Greg Hubert, B.S.
Elaine Keller, M.Ed.
Kathryn M. McDonald, M.M.
Patrick Romano, M.D., M.P.H
Jeff Schoenborn, B.S.

The **Agency for Healthcare Research and Quality** Support for Quality Indicators-II team includes:

Marybeth Farquhar, Project Officer
Mamatha Pancholi, Project Officer
Mary B. Haines, Contract Officer

This product is based on the work of many individuals who contributed to its development and testing.

The following staff from the **Evidence-based Practice Center (EPC) at UCSF-Stanford** performed the evidence review, facilitated clinician reviews, completed the empirical evaluation, and created the programming code and technical documentation for the Prevention Quality Indicators.

Core Project Team
Mark McClellan, M.D., Ph.D., *principal investigator*
Kathryn M. McDonald, M.M., *EPC coordinator*
Sheryl M. Davies, M.A.
Jeffrey Geppert, J.D.
Patrick Romano, M.D., M.P.H.
Kaveh G. Shojania, M.D.

Other Contributors
Amber Barnato, M.D.
Paul Collins, B.A.
Bradford Duncan M.D.
Michael Gould, M.D., M.S.
Paul Heidenreich, M.D.
Corinna Haberland, M.D.
Paul Matz, M.D.
Courtney Maclean, B.A.
Susana Martins, M.D.
Kristine McCoy, M.P.H.
Suzanne Olson, M.A.
L. LaShawndra Pace, B.A.
Mark Schleinitz, M.D.
Herb Szeto, M.D.
Carol Vorhaus, M.B.A.
Peter Weiss, M.D.
Meghan Wheat, B.A.

Consultants
Douglas Staiger, Ph.D.

The following staff from **Social & Scientific Systems**, **Inc.**, developed the original software product, documentation, and guide:

Programmers
Leif Karell
Kathy McMillan
Fred Rohde

Technical Writer
Patricia Caldwell

Graphics Designer
Laura Spofford

Contributors from the **Agency for Healthcare Research and Quality**:

Anne Elixhauser, Ph.D.
H. Joanna Jiang, Ph.D.
Denise Remus, Ph.D., R.N.
Margaret Coopey, R.N., M.G.A, M.P.S.

We also wish to acknowledge the contribution of the peer reviewers of the evidence report and the beta-testers of the software products, whose input was invaluable.

Table of Contents

Abstract ..... **ii**

Acknowledgments ..... **iii**

1.0 Introduction **1**

2.0 Components of the Prevention Quality Indicators Module **2**

3.0 Quick Reference **2**

4.0 Background **6**

5.0 Data Elements and Coding Conventions **7**
- 5.1 Input for the AHRQ Quality Indicators 7
- 5.2 Coding of Diagnoses, Procedures, and DRGs 9
- 5.3 Missing Values 12
- 5.4 Treatment of Missing Data 12

6.0 User Options **13**

7.0 Processing Steps **13**
- 7.1 Step 1: Identify Outcomes in Inpatient Records 14
- 7.2 Step 2: Identify Populations at Risk from Census Population Data 14
- 7.3 Step 3: Calculate Observed (Raw) Prevention Quality Indicator Rates 14
- 7.4 Step 4: Risk Adjust the Prevention Quality Indicator Rates 14
- 7.5 Step 5: Create MSX Smoothed Rates 15
- 7.6 Step 6: Calculate Condition-specific Rates 15

8.0 Program Descriptions **15**
- 8.1 Program 1: CONTROL_PQI.SAS 15
- 8.2 Program 2: PQFMTS.SAS 16
- 8.3 Program 3: PQSASA1.SAS 17
- 8.4 Program 4: PQSASA2.SAS 18
- 8.5 Program 5: PQSASA3.SAS 20
- 8.6 Program 6: PQSASC2.SAS 22

9.0 Reviewing the Printed Output **23**
- 9.1 PQSASA1.SAS 23
- 9.2 PQSASA2.SAS 24
- 9.3 PQSASA3.SAS 24
- 9.4 PQSASC2.SAS 25

10.0 Benchmark Timings **27**

11.0 User Support **27**

Appendix A: Links **A-1**

List of Tables

Table 1. Prevention Quality Indicator (PQI) Variables 5
Table 2. Prefixes for the Prevention Quality Indicator (PQI) Variables 5
Table 3. Prevention Quality Indicator Module Contents 6
Table 4. Data Elements and Coding Conventions 10
Table 5. Treatment of Missing Data Elements 13
Table 6. PQSASA2 aggregation possibilities 19
Table 7. Area Level Printed Output 25

List of Figures

Figure 1. Processing Steps for the PQIs - Area 3
Figure 2. Processing Steps for the PQIs - Condition-specific 4

AHRQ Quality Indicators, Software Documentation, Version 3.0a: Prevention Quality Indicators

1.0 Introduction

This documentation describes the software for implementing the Prevention Quality Indicator (PQI) module of the AHRQ Quality Indicators Version 3.0a. Detailed information about the Prevention Quality Indicators (PQI) is contained in the companion document *Guide to the Prevention Quality Indicators,* which includes information about the development of the indicators and their definitions. Detailed definitions, including all ICD-9-CM codes included in or excluded from the numerator and denominator for each indicator, are contained in a separate document *Prevention Quality Indicators Technical Specifications.* See Appendix A for links to these and other documents as well as Web sites that may be of interest to PQI users.

The section called "Interpreting the Results" has been removed from this document. The *Guide to the Prevention Quality Indicators* contains a new section, "Using Different Types of QI Rates."

The software consists of a series of computer programs that:

- Assign and calculate PQIs from hospital discharge abstracts or UB (uniform billing) claims data.
- Print PQI results in SAS format or as an ASCII text file (comma delimited) that can be imported into word processing documents, spreadsheets, or graphics software, at the user's option.
- Provide four options for stratifying rates by geographic areas.
- Calculate rates using either the hospital location or the county of patient residence.
- Create risk-adjusted rates that adjust for casemix differences.
- Create smoothed rates that reduce fluctuations over time due to random variation.
- Provides the option to generate condition-specific rates (diabetes) by state and age.

This documentation includes:

- A discussion of the data elements necessary to produce the AHRQ PQIs and the uniform coding conventions recommended for those elements.
- Descriptions of the PQI processing steps in nontechnical language.
- Detailed descriptions of the functions of each PQI SAS program.

Critical user information is highlighted with this symbol:

2.0 Components of the Prevention Quality Indicators Module

As shown in Table 3 on page 6 the Prevention Quality Indicators module consists of six SAS programs and 13 ASCII text files that contain auxiliary data. These programs and text files are described in the subsequent sections of this document. The programs were developed for use in Statistical Analysis System (SAS)[1], version 6.12 or higher, on a Windows personal computer.

A note on the naming conventions for the Prevention Quality Indicators (PQI) module programs. The programs have names of the form PQSASAi. The first two characters "PQ" of the program name indicate a Prevention Quality Indicators program, to distinguish it from other modules that have been or will be released from AHRQ. The next three characters of the program name are "SAS" and are present to distinguish the program from SPSS versions of the software. The sixth character of the program name is either "A" to denote a program that is for the production of Area indicator rates that use county or Office of Management and Budget (OMB) defined metropolitan or micropolitan area (Metro Area) populations as denominators, or "C" to denote a program for the production of condition-specific indicator rates[2]. The last character (i) designates the number of the specific program.

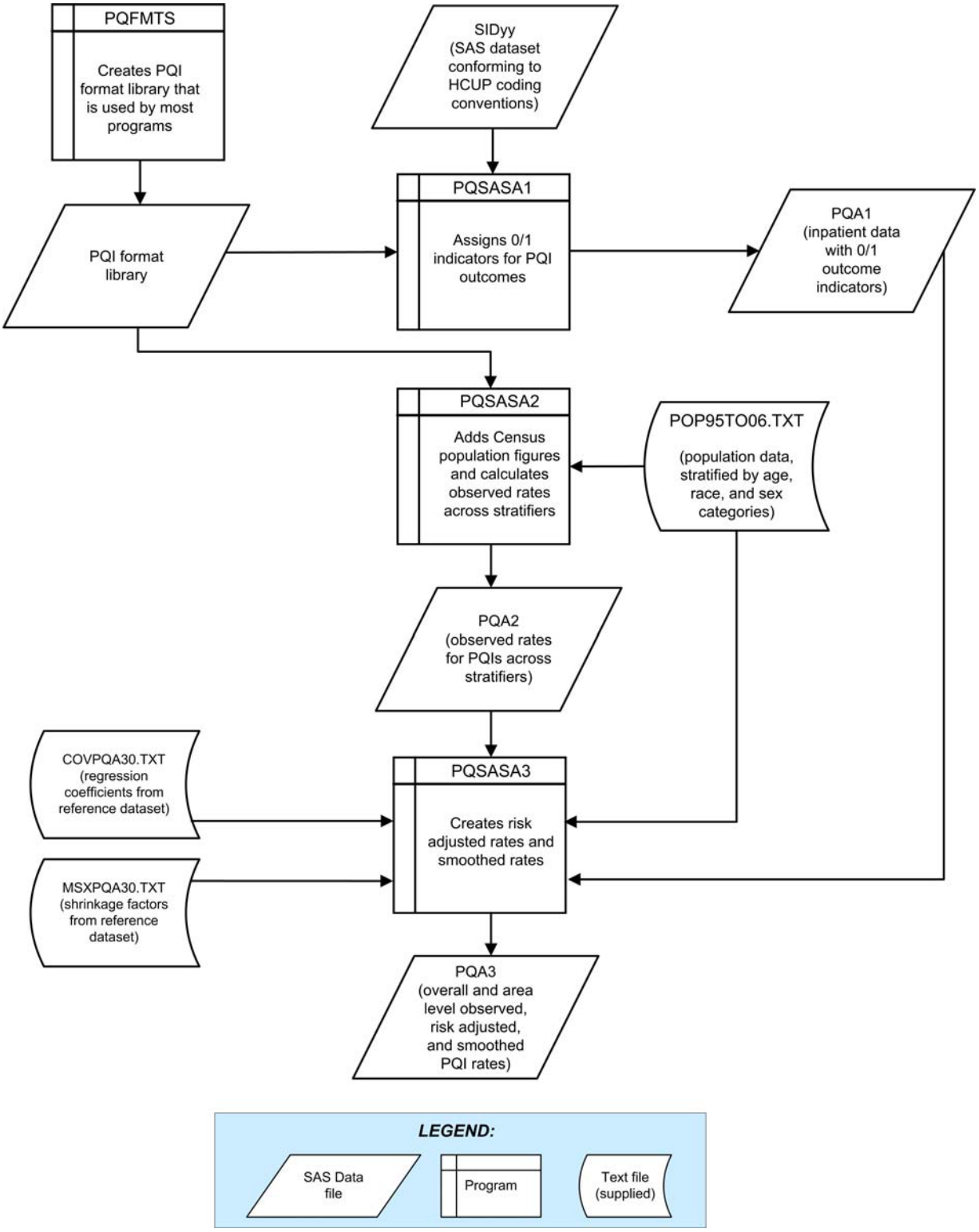
3.0 Quick Reference

The subsequent three pages provide information intended to serve as a quick reference to assist in reading this documentation and in reviewing the Prevention Quality Indicator (PQI) module outputs. Processing steps (Figure 1 and Figure 2) are shown first followed by a listing of the module variables (Table 1), followed by variable prefixes (Table 2), and finally a listing of the SAS module contents including the SAS syntax files and text reference files used by the program (Table 3).

[1] SAS is a statistical program distributed by the SAS Institute, Inc. (http://www.sas.com). The company may be contacted directly regarding the licensing of their products. The SAS Institute, Inc. does not have any affiliation with AHRQ nor involvement in the development of the AHRQ QIs.

[2] The software provides the ability to produce condition-specific indicator rates for the four diabetes PQIs (PQIs 1, 3, 14, and 16). The user can calculate these PQIs based on the area population or by the prevalence of condition, i.e., the number of diabetics in a State stratified by age.

Figure 1. Processing Steps for the PQIs - Area



NOTE: Each of the SAS Programs in Figure 1 incorporates the CONTROL_PQI.SAS program, which contains code the user **must** modify in order to run the PQI software. The **modifications** include such items as specifying the name and location of the input data set, the population data set, and the output data sets.

Figure 2. Processing Steps for the PQIs - Condition-specific

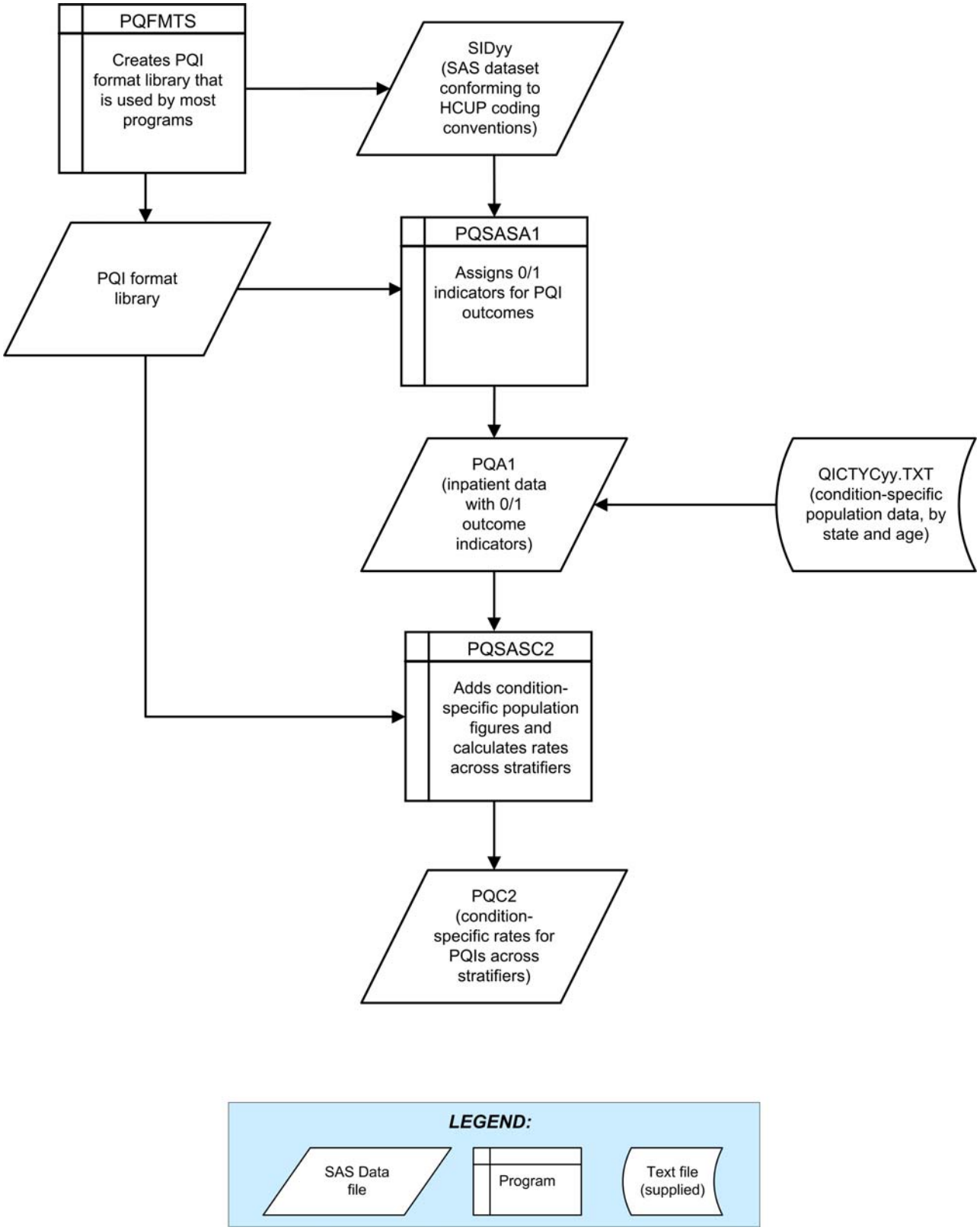


Table 1 lists the numerators and denominators for each of the PQIs.

Table 1. Prevention Quality Indicator (PQI) Variables

Indicator number	Numerator (admissions for ACSC)	Denominator
1	Diabetes short-term complication	Area population
2	Perforated appendix	**Appendicitis***
3	Diabetes long-term complication	Area population
5	Chronic obstructive pulmonary disease	Area population
7	Hypertension	Area population
8	Congestive heart failure	Area population
9	Low birth weight	Births
10	Dehydration	Area population
11	Bacterial pneumonia	Area population
12	Urinary tract infection	Area population
13	Angina without procedure	Area population
14	Uncontrolled diabetes	Area population
15	Adult asthma	Area population
16	Lower-extremity amputation among patients with diabetes	Area population

ACSC – Ambulatory Care Sensitive Condition
* The denominator includes all hospital admissions for this condition in the area.

The population rate is based on all discharges in the 2003 HCUP State Inpatient Databases for 38 states. Except for PQI 2, population rates are expressed per 100,000.

Two of the original PQIs have been moved to the Pediatric Quality Indicators module, and four have been revised to exclude patients under the age of 18. All PQIs now apply only to adult populations.

Table 2 lists the prefixes used for used for PQI variable names.

Table 2. Prefixes for the Prevention Quality Indicator (PQI) Variables

Prefix	Contents	Example (for PQI # 1, Diabetes short–term complications)
T	Inpatient numerator (top)	TAPQ01
P	Population denominator (pop)	PAPQ01
O	Observed Rate	OAPQ01
E	Expected Rate	EAPQ01
R	Risk-adjusted Rate	RAPQ01
L	Lower limit of confidence interval of risk-adjusted rate	LAPQ01
U	Upper limit of confidence interval of risk-adjusted rate	UAPQ01
S	Smoothed Rate	SAPQ01
X	Standard error of smoothed rate	XAPQ01
C	Condition-specific Rate	CAPQ01

Table 3. Prevention Quality Indicator Module Contents

SAS programs (syntax files)	ASCII text files (data)	
CONTROL_PQI.SAS	POP95TO06.TXT	QICTYC03.TXT
PQFMTS.SAS	QICTYC97.TXT	QICTYC04.TXT
PQSASA1.SAS	QICTYC98.TXT	QICTYC05.TXT
PQSASA2.SAS	QICTYC99.TXT	QICTYC06.TXT
PQSASA3.SAS	QICTY**C**00.TXT	
PQSASC2.SAS	QICTYC01.TXT	COVPQA30.TXT
	QICTYC02.TXT	MSXPQA30.TXT

The ASCII text files are provided with the Prevention Quality Indicators module, and are necessary for the programs to run. There is no need to convert these files to SAS format for use with the software. The parameter reference files of MSXPQA30.TXT and MSXPQA_3.TXT are not interchangeable between SAS and SPSS software. Users should verify they are running the software with the ASCII text files as listed in Table 3 of the applicable, e.g., SAS or SPSS, software manuals.

4.0 Background

The PQI module contains 14 indicators that measure adult hospital admissions for ambulatory care sensitive conditions (ACSC) across geographic areas. ACSCs represent conditions for which hospitalization could be avoided if the patient receives timely and adequate outpatient care. Many factors influence the quality of outpatient care, including access to care and adequately prescribed treatments, once care is obtained. In addition, patient compliance with those treatments and other patient factors may play a role. In general, areas with lower social-economic status tend to have higher admission rates for ACSCs than areas with higher social-economic status. As with utilization indicators, there are no “right rates” of admission for these conditions. Very low rates could signal inappropriate underutilization of healthcare resources while very high rates could indicate potential overuse of inpatient care.

Therefore, hospital admission for ACSCs is not a measure of hospital quality but a potential indicator of outpatient and community health care need. For example, if an area has a relatively high hospital admission rate for diabetes complications, the local healthcare providers should work with the community to identify reasons and strategies to address the problem.

The 14 ACSCs in the PQI module are listed in Table 1 on page 5. See also *Prevention Quality Indicators Technical Specifications* for the definition of each indicator.

Except for perforated appendix, each indicator is measured as the number of hospital admissions for a particular ACSC per 100,000 residential population in an area.

Geographic areas can be defined at the user's option in one of four ways:

- County level with U.S. Census FIPS[3]
- County level with modified FIPS
- Metro Area with OMB 1999 definition
- Metro Area with OMB 2003 definition.

The MSALEVL parameter has been renamed to MALEVL to reflect the change in OMB definitions for areas. Refer to page 8 for information on these area definitions and MALEVL parameter settings.

[3] Federal Information Processing Standard, as defined by the U.S. Department of Commerce, National Institute of Standards and Technology (formerly National Bureau of Standards).

The data required for measuring these indicators come from hospital discharge abstracts or billing claims (administrative data) which are readily available in many states. The residential population data are from the U.S. Census Bureau.

The software generates observed, risk-adjusted, and smoothed rates for each indicator at the area level. Observed rates are the raw rates. Risk-adjusted and expected rates are derived from applying the average casemix of a baseline file that reflects a large proportion of the U.S. hospitalized or residential population. Smoothed rates are estimates with removal of fluctuations over time due to random variation.

For more detail and information about how these indicators were identified, see the *Guide to the Prevention Quality Indicators.*

5.0 Data Elements and Coding Conventions

5.1 Input for the AHRQ Quality Indicators

The PQI software was written to process data from discharge data abstracts that contain information about hospital stays. The specific data elements that are needed for the PQI software are listed in Table 4. The PQI module was tested on data from the Healthcare Cost and Utilization Project (HCUP). HCUP is an ongoing Federal-State-private collaboration to build uniform databases from administrative hospital-based data collected by State data organizations and hospital associations.

The input data files for the Prevention Quality Indicators software must be in SAS.

We recommend that users recode data elements in their input files to be consistent with the coding expected by the software. This will minimize the number of internal changes that will be necessary in the Prevention Quality Indicators software. For example, recoding the SEX data element in the input file to be consistent with the coding described in Table 4 (i.e., 1 for male and 2 for female) is easier than modifying all uses of the SEX data element in the PQI programs.

Not every value for every data element is referenced by the PQI software. For example, admission source (ASOURCE) values are used only to identify transfers. However, we include the complete set of HCUP values to assist users of the uniform HCUP data files.

To minimize internal changes to the software, all required elements should be present in the input data file. If a required element is not available, a dummy element should be provided. Failure to provide a dummy element will result in errors during the execution of the PQI programs.

The data elements listed in Table 4 are those required for the input data files of **all** the SAS and SPSS QI modules. When a variable is not required for the PQI SAS program the notation “Not used by the PQI program” will be present in the Comments column. Standardizing the variables and data values in the input data file will be helpful for users who have access to both statistical software packages or want to set up their input data file for use with any of the three QI modules: Inpatient Quality Indicators (IQIs), Prevention Quality Indicators (PQIs) or the Patient Safety Indicators (PSIs).

Table 4 has been standardized to represent the input data file specifications required for use with any of the three QI modules (IQI, PQI or PSI). If a specific variable is not used by the PQI program it is not necessary to create a dummy variable in the input data file. For example, the variable APR-DRG is used only by the IQI software so it is not necessary to have this data element in the input data file or create a dummy variable for the PQI software program to run correctly.

The data element HOSPSTCO has been renamed to PSTCO to emphasize the importance of calculating PQIs by the location of the patient residence. If the user wants to calculate the PQIs based on the

population of the Metro Area or county associated with the patient residence, the values for this variable should be the FIPS state/county code associated with each individual patient's place of residence. If the patient information is not available or the user wants to calculate the PSIs using the population associated with the hospital location as the denominator, the values for this variable should be the individual hospital FIPS state/county codes.

If the hospital FIPS code is used in PSTCO, users should be aware that rates may be biased for hospitals which serve as regional referral centers. These hospitals are likely to treat patients from outside the Metro Area, county or even the state in which the facility is located.

Users no longer have to use the Modified FIPS codes assignment for area denominators. However, that option is still available. In the modified FIPS codes certain independent cities (Baltimore City, Carson City and St. Louis City), and areas within Hawaii and Virginia, are assigned to different area groupings in the Modified FIPS categories.

Using the patient FIPS state/county code for analysis may more accurately reflect the true population at risk. Evaluation of geographic variations in admissions for ambulatory care sensitive conditions by patient Metro Area or county of residence can result in better information to guide community or provider response. Using the hospital FIPS state/county code for analysis should be done with caution, and at larger levels of geographic area (i.e. Metro Area) to minimize the bias with patients that come from a different county than the hospital location.

It is possible that some records in the input data file may be missing the patient FIPS code. Users should be aware that any records with missing values (in the PSTCO data field) are excluded from the calculations of observed, risk-adjusted and smoothed PQI rates. They will be included in the output from the third program (PQSAS1.SAS.)

The software will generate rates for each county included in the PSTCO data field. Users are encouraged to limit cases in their input file to those patients discharged from the geographic area of interest. For example, if you are using data from the state of Wisconsin and are interested in generating a rate for Wisconsin, you should remove the cases where the patient's county of residence (FIPS code) is from another state. Otherwise the total or overall rate will include the population of the counties outside Wisconsin.

This version of the software provides the user with the option of producing output by Metro Area or by county. When Metro Area is selected, urban areas are defined by Metro Areas. When county is selected urban areas will be defined by county. Rural areas are always defined by county.

In the software programs, the MALEVL parameter should be set as follows:

0-County level with U.S. Census FIPS
1-County level with modified FIPS
2-Metro Area with OMB 1999 definition
3-Metro Area with OMB 2003 definition

The metropolitan area definitions are from three different sources. 1) The "modified FIPS" definition is from the Area Resource File. The mapping is from county to modified FIPS county (e.g., Baltimore city to Baltimore county); 2) The "1999 OMB" definition is from the federal Office of Management and Budget (OMB) circular 99-04, (Last Revised May 6, 2002). The mapping is from county to Metropolitan Statistical Area (MSA), except in New England where counties are assigned to NECMAs (New England County Metropolitan Area). OMB defines NECMAs as a county-based alternative to the city- and town-based New England MSAs and CMSAs (Consolidated MSAs). For example, Hampden and Hampshire counties in Western Massachusetts are assigned to the Springfield MA NECMA, even though Holland town in Hampden County is part of the Boston MSA; and 3) the "2003 OMB" definition is from the federal Office of

Management and Budget (OMB) circular 03-04, (Last Revised December 4, 2005). The mapping is from county to either Metropolitan Statistical Area or Micropolitan Statistical Area.

5.2 Coding of Diagnoses, Procedures, and DRGs

Diagnoses and procedures must be coded using the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM)

Significant modifications to ICD-9-CM codes occurred in the early 1990s. PQI definitions only reflect codes valid after October 1, 1994, and therefore may not accurately analyze data collected before 1994.

- Diagnoses and procedure codes should be in character strings.
- Diagnosis and procedure codes should not contain any decimal points.
- Diagnosis and procedure codes should be left justified. Diagnosis codes with fewer than five digits should be padded with spaces (not zeros) on the right. Procedure codes with fewer than four digits should be padded with spaces (not zeros) on the right.

The PQI programs assume accurate and specific coding of ICD codes. If an ICD code is specified using 5 digits, the software recognizes only 5-digit versions of that code and not versions truncated at 3 or 4 digits. Similarly, if an ICD code is specified using 4 digits, the software recognizes on 4-digit versions of that code and not versions truncated at 3 digits. For example, 4281 is the ICD-9-CM diagnosis code for left heart failure. If your data are coded less specifically (i.e., 428 is used to designate "heart failure"), these less specific codes are not recognized by the software and will result in undercounting cases.

- This software has ICD-9-CM codes updated through FY 2006 changes.
- Diagnosis-related groups (DRGs) are those derived from the Centers for Medicare & Medicaid Services (previously Health Care Financing Administration) Medicare grouper. The software expects that you will be using the DRG effective on the discharge date. The software now refers to data elements DRG and MDC. Your data should be coded accordingly. The software may be modified at the user's option to use other types of DRGs. However, the impact of using other types of DRGs should be evaluated carefully before making such a change.

Version 3.0a of the software accounts for ICD-9-CM and DRG coding effective through FY 2006 (or through September 30, 2006).

Table 4. Data Elements and Coding Conventions

Variable name	Description	Format	Value description	Comments
KEY	Sequence number. Unique case identifier.	Numeric	User defined unique numeric identifier for each discharge record	Not used by the AHRQ QI programs, but should be present to facilitate possible exploration; allows user to link the records from the PQSAS1.SAS program output file back to the original input data file.
AGE	Age in years at admission	Numeric	Age in years	If this data element is missing the discharge record will be excluded from analysis.
AGEDAY	Age in days at admission (coded only when the age in years is less than 1)	Numeric	Age in days	Used for PQI #9, PSI #17, and all of the PDIs. If this data element is missing, the value is treated as if it was less than 30 days.
RACE	Race of patient.	Numeric 1 2 3 4 5 6	 White Black Hispanic Asian or Pacific Island Native American Other	The values 1through 5 are used directly in the QI software to stratify rates. All other ethnicity codes are mapped to an 'other' category.
SEX	Gender of patient	Numeric 1 2	 Male Female	If this data element is missing the discharge record will be excluded from analysis.
PAY1	Expected primary payer	Numeric 1 2 3 4 5 6	 Medicare Medicaid Private, incl. HMO Self-pay No charge Other	Not used by the PQI program. If not present in the input data file it is **not** necessary to create a dummy variable.
PSTCO	Location of patient residence or hospital location (FIPS† State/county code)	Numeric ssccc	Modified Federal Information Processing Standards State/County code	See Appendix A for link to list of most recent codes. If this data element is missing the discharge record will be excluded from rate calculations.
HOSPID	Data Source hospital number	Numeric	Hospital identification number	Used to facilitate data exploration and possible trouble shooting.
DISP	Disposition of patient	Numeric 1 2 3 4 5 6 7 20	 Routine Short-term hospital Skilled nursing facility Intermediate care Another type of facility Home health care Against medical advice Died	Not used by the PQI program. If not present in the input data file it is **not** necessary to create a dummy variable.
ATYPE	Admission Type	Numeric 1 2 3 4 5 6	 Emergency Urgent Elective Newborn Trauma Center[4] Other	Value of 4 used to identify live births

[4] The ATYPE value "5" was used to indicate delivery only in the 1988 to 1997 HCUP data files. The UB standards currently use "5" to indicate Trauma Center.

Variable name	Description	Format	Value description	Comments
ASOURCE	Admission Source	Numeric 1 2 3 4 5	ER Another hospital Another fac. incl. LTC Court/law enforcement Routine/birth/other	The values 2 and 3 are referenced by the PQI code (to identify transfers from another hospital)
LOS	Length of Stay	Numeric	Number of days from admission to discharge	Not used by the PQI program. If not present in the input data file it is **not** necessary to create a dummy variable.
APR_DRG	APR-DRG category	Numeric	APR-DRG from 3M software.	Not used by the PQI program. If not present in the input data file it is **not** necessary to create a dummy variable.
SEVERTY	APR-DRG Severity Score	Numeric	APR-DRG Severity Score from 3M software.	Not used by the PQI program. If not present in the input data file it is **not** necessary to create a dummy variable.
RISKMORT	APR-DRG Mortality Score	Numeric	APR-DRG Risk of Mortality Score from 3M Software.	Not used by the PQI program. If not present in the input data file it is **not** necessary to create a dummy variable.
DRG	Diagnosis Related Group	Numeric	DRG from federal (CMS) DRG Grouper	If this data element is missing the discharge record will be excluded from analysis.
MDC	Major Diagnostic Category	Numeric	MDC from federal (CMS) DRG grouper	If this data element is missing the discharge record will be excluded from analysis.
DX1 – DX30	Diagnoses ICD-9-CM codes. DX1 is the principal diagnosis, DX2-DX30 are secondary diagnoses.	String, 5 characters, left-justified (e.g. 020.2 -> 0202_ where _ indicates a space)	Diagnosis codes	Users with more or fewer secondary diagnoses should modify the parameter &NDX in CONTROL_PQI.SAS to reflect the number of diagnoses.
NDX	Count of the number of diagnoses (Dx) on the discharge record.	Numeric	Count of diagnoses (principal and all secondary diagnoses)	If the 1st discharge record has 5 Dx and the 2nd has 10, NDX would be 5 on the 1st and 10 on the 2nd. Not used by the SAS program.
PR1 – PR30	Procedure ICD-9-CM codes. PR1 is the principal diagnosis, PR2-PR30 are secondary procedures.	String, 4 characters left-justified	Procedure code	Users with more or fewer secondary procedures should modify the parameter &NPR in CONTROL_PQI.SAS to reflect the number of diagnoses.
NPR	Count of the number of procedures (Pr) on the discharge record.	Numeric	Count of procedures (principal and all secondary procedures)	If the 1st discharge record has 1 Pr and the 2nd has 2, NPR would be 1 for the 1st record and 2 for the 2nd. Not used by the SAS program.
PRDAY1 – PRDAY30	Days from admission to procedure. PR1 is the principal procedure, PR2-PR30 are secondary procedures.	Numeric	Days from Admission to Procedure	Not used by the PQI program. If not present in the input data file it is **not** necessary to create a dummy variable.

Variable name	Description	Format	Value description	Comments
YEAR	Year of discharge. The patient's year of discharge. For example, a patient discharged on July 7, 2004 would have a discharge year of '2004.'	Numeric	YYYY Discharge year should be within the range of 1997 to 2006.	
DQTR	Quarter of discharge. The calendar quarter of the patient's discharge. For example, a patient discharged on July 7, 2004 would have a discharge quarter of '3.'	Numeric	1 January-March 2 April-June 3 July-September 4 October-December	
	AHRQ comorbidity variables*			Required if Risk-Adjusted rates are desired.

†Federal Information Processing Standard, as defined by the U.S. Department of Commerce, National Institute of Standards and Technology (formerly National Bureau of Standards).

Note: Certain independent cities (Baltimore City, Carson City and St. Louis City), and areas within Hawaii and Virginia, are assigned to different area groupings in the Modified FIPS categories as compared to the US Census Dept. groupings. The AHRQ QI software uses the Modified FIPS code assignment of these areas. Failure to use the Modified FIPS codes assignment will result in the use of incorrect denominator estimates for area indicators.

5.3 Missing Values

The PQI programs do not distinguish among different types of missing values.

Cases that are missing data in fields used for indicator definitions or for risk adjustment (such as age and sex) will be excluded from analysis.

Data files of hospital discharge abstract data may have numeric data elements coded using special SAS missing "dot" values. For example, besides the standard SAS value of "." for missing data, there might also be present values of ".A" for invalid data, ".B" for data unavailable from a particular source, ".C" for inconsistent data. However, the PQI programs do **not** distinguish among the different types of missing codes. Therefore, all types of missing values in the input data to the PQI module can be represented by a single value - missing data (.) for numeric variables and blank (' ') for alphanumeric (or character) variables.

5.4 Treatment of Missing Data

The software is designed to handle missing data in a specific manner; specifically, the software requires confirmation for the assignment of a poor outcome or negative event. For instance, in order to be excluded as a transfer, each case must actually be coded as a transfer. Missing data is considered neutral.

Missing data for some elements results in the exclusion of that case from the numerator or denominator. For a few other elements, the case is retained. Table 5 lists the impact of missing data for each data element.

Table 5. Treatment of Missing Data Elements

Variable	Label	D	N	E	S	Treatment of Missing Data
AGE	Age In Years At Admission	X	X		X	Case excluded from all analysis (e.g., dropped from numerator of all population based measures and from the denominator of the discharge based indicator, PQI 2.
ASOURCE	Admission Source			X		Transfers from another institution cannot be identified and therefore will not be excluded from the denominator.
DRG	Diagnosis Related Group (CMS DRG) in Effect on Discharge Date	X	X			Case excluded from all analysis (e.g., dropped from numerator of all population based measures and from the denominator of the discharge based indicator, PQI 2.
SEX	Patient Gender	X	X		X	Case excluded from all analysis (e.g., dropped from numerator of all population based measures and from the denominator of the discharge based indicator, PQI 2.
PSTCO	Location of Patient Residence or Hospital Location Modified FIPS State/County Code	X			X	Dropped from denominator in rate calculations (stratified by county or Metro Area), case will appear in calculation of the overall rate.
MDC	Major Diagnostic Category (CMS MDC) In Effect On Discharge Date	X	X			Case excluded from all analysis (e.g., dropped from numerator of all population based measures and from the denominator of the discharge based indicator, PQI 2.
RACE	Race				X	Classified As “Other”

D – Denominator; N – Numerator; E – Exclusion; S – Stratification

6.0 User Options

The PQI software reflects the development and implementation of the software in data available to AHRQ. Our goal was to develop the tools, illustrate their use, and then encourage others to adopt and use the tools for their own applications. As a result, we expect and encourage users of the software to consider whether and how to modify the PQI software to better serve their local conditions or interests.

Modifications to the definitions of outcomes of interest (numerators) or populations at risk (denominators) are possible but not desirable. Maintaining consistent definitions is important. Once definitions are altered, the ability to make comparisons of PQIs based on original definitions is lost. We encourage users to identify ways to improve the PQI methodology and to share their suggestions with us for future updates.

AHRQ has provided national estimates using the PQIs through its web-based query system, HCUPnet (http://hcup.ahrq.gov/HCUPnet.asp). If users change definitions of the PQIs, it will not be possible to compare users’ results to the national estimates in HCUPnet.

7.0 Processing Steps

Each Prevention Quality Indicator (PQI) expressed as a rate, is simply defined as:

Outcome of interest / Population at risk

Conceptually, six steps are necessary to produce the PQI rates. The following describes the steps and how the software performs them.

7.1 Step 1: Identify Outcomes in Inpatient Records

Inpatient records are marked to indicate whether they contain the outcome of interest (numerator or "top") for each of the AHRQ PQI measures.

This is done by setting a series of flag variables, each of which corresponds to the numerator for a particular PQI. For example, if the inpatient record meets the conditions for inclusion in the outcome for PQI #1, then the outcome indicator for PQI #1 is set to 1 for that record.

This step requires one pass through the discharge-level data and outputs a discharge-level data file containing the flag indicator variables for the outcomes for each PQI. This output file also contains the original input variables, with the exception of the diagnosis and procedure variables which have been removed to reduce the size of the output file.

7.2 Step 2: Identify Populations at Risk from Census Population Data

The populations at risk (the denominators for calculating the PQI rates) are derived from Census population figures by areas defined by county in all areas of the US, or by Metro Area in urban areas and by county in rural areas.

7.3 Step 3: Calculate Observed (Raw) Prevention Quality Indicator Rates

Using the output data from step 1 and Census population data from step 2, the PQI rates are calculated for user-specified combinations of stratifiers. These stratifiers include area, age groups, race/ethnicity categories, and sex.

The programs calculate observed PQI rates regardless of the number of cases available. However, PQI rates based on only a few cases should be interpreted with caution.

In some performance measurement work, it is often recommended that rates be suppressed when there are fewer than 30 cases in the denominator. This exclusion rule serves two purposes:

- It eliminates unstable estimates based on too few cases.
- It helps protect the identities of hospitals and patients.

This step outputs the observed rates to user-specified levels by area, age, sex, and race/ethnicity categories, alone or in combination. For example, rates can be output at the area level, at the age group level, or at the area level by age groups.

7.4 Step 4: Risk Adjust the Prevention Quality Indicator Rates

Overall file means and regression coefficients from a baseline database (reflecting a large proportion of the U.S. population) are applied to the observed rates to risk-adjust the rates observed in the user's data. These baseline file means and regression coefficients are provided as part of the PQI module. The risk-adjusted rates will then reflect the age and sex distribution of areas in the baseline file rather than the distribution for the areas covered by the users' data. This will allow risk-adjusted rates produced by various users to be compared directly to one another. The overall means and regression coefficients were derived from AHRQ's State Inpatient Databases (SID) for

38 states. The codes to generate these baseline means and coefficients are not part of the PQI module.

7.5 Step 5: Create MSX Smoothed Rates

Shrinkage factors are applied to the risk-adjusted rates for each PQI in a process called multivariate signal extraction (MSX). These shrinkage factors were calculated from the SID database of 38 states. For each PQI, the shrinkage estimate reflects a 'reliability adjustment' unique to each indicator. The less reliable the PQI over time and across areas, the more the estimate 'shrinks' the PQI toward the overall area mean. The resulting rate will appear "smoother" than the raw rate, meaning the year-to-year fluctuations in performance are likely to be reduced. The shrinkage factors are provided as part of the PQI software and do not need to be calculated by users.

7.6 Step 6: Calculate Condition-specific Rates

For selected indicators, data are available on the prevalence of the relevant condition. Version 3.0a of the software incorporates state level estimates of diabetes prevalence by age from the CDC National Diabetes Surveillance System (see Appendix A.) The four PQIs related to diabetes (PQI #1, 3, 14 and 16) can be calculated using the number of diabetics in the state as the denominator, stratified by age.

These six steps reflect the PQI module production in a nutshell. The next section of this document describes the specifics of each software component of the PQI module software.

8.0 Program Descriptions

This section describes the six SAS programs that assign, calculate, and print the Prevention Quality Indicators.

For each program there is a description, a list of input and output files, and an explanation of changes to the program that may be required. The flow of data through the PQI module programs is shown in the flowcharts in Figure 1 and Figure 2 that begin on page 3.

If you desire to create and examine observed PQI rates, then the PQFMTS, CONTROL_PQI, PQSASA1, and PQSASA2 programs will need to be used. If you also wish to create risk-adjusted and smoothed PQI rates, then you will also need to run the PQSASA3 program. If you wish to create condition-specific (diabetes) rates, you will need to run the PQSASC2 program.

8.1 Program 1: CONTROL_PQI.SAS

The CONTROL_PQI.SAS code contains all of the SAS statements that you need to modify in order to run the remaining PQI module programs (with one exception noted below). This is where you specify, for example, the path names or locations of all your input and output data sets and whether you want rates calculated at the county level or the Metro Area/county level. There is also a TITLE line that the user can fill in. The items that need to be changed in order to run the PQSASA1, PQSASA2, and PQSASA3 programs are described below under each program.

*The CONTROL_PQI.SAS code is automatically inserted into the remaining programs when they are executed. You do **not** need to run CONTROL_PQI.SAS as a separate program.*

All of the changes needed for the CONTROL_PQI.SAS code are listed in the following descriptions for programs 2 – 5. Each program description contains a section labeled "Changes:" that enumerates the modifications for CONTROL_PQI.SAS that are associated with that particular program.

*Lines of code to be modified are identified by the phrase "<===USER MUST modify" which means the user **must** make the change or by "<===USER MAY modify" which means the modification is optional.*

In all of the remaining programs included with the PQI module, there is one line of code near the beginning of each program that needs to be modified. This line of code starts with the phrase "FILENAME CONTROL". You ***must*** modify this line to point to the path name where you stored your modified copy of the CONTROL_PQI.SAS file.

CONTROL_PQI.SAS has a collection of parameters that allow you to read in data from and write out data to different locations or pathnames. For example, "LIBNAME1 IN1" points to the location of your input data set for program PQSASA1 and "LIBNAME OUT1" points to the location of the output data set created by the PQSASA1 program. The parameters allow these locations to be different from one another. However, if you wish to read in and write out all of the data to the same location, it is easiest to just make a global change in CONTROL_PQI.SAS changing "C:\pathname\" to the single directory location that you wish to use.

There is a global PRINT parameter in CONTROL_PQI.SAS that as a default prints all of the output summary records at the end of the PQSASA2 and PQSASA3 programs. These printouts may be quite large if you have a large database with a large number of areas (eg, counties, Metro Areas) and if you choose the results to be finely stratified. If you wish to turn these prints on, set PRINT to have a value of 1 in CONTROL_PQI.SAS.

The global MALEVL parameter in CONTROL_PQI.SAS by default instructs all programs to define geographic areas by county for all areas in the input datasets. Refer to Page 8 for a list of values for the MALEVL parameter.

8.2 Program 2: PQFMTS.SAS

The PQFMTS program defines a format library, which contains the diagnosis and procedure screens necessary for assigning the outcomes of interest. This format library is used by most of the succeeding programs.

The first few formats in this program are for the stratifiers (age, sex, race/ethnicity). These are followed by formats that are used for the identification of outcomes of interest in the PQSASA1 program, and by formats that map FIPS county codes to Metro Areas in the PQSASA1 and PQSASA2 programs. At the end of the program are formats used to assign age categories for the risk adjustment process implemented in PQSASA3.

The PQI module Version 3.0a contains formats that incorporate ICD-9-CM and DRG coding through September 30, 2006.

Input: None.

Output: Permanent SAS format library (FMTLIB).

Changes:

1. You must modify the "FILENAME CONTROL" statement at the beginning of the program to change the path so that it points to your location of the CONTROL_PQI.SAS code.

2. In CONTROL_PQI.SAS, specify the location (path) of the format library in the "LIBNAME LIBRARY" statement.

8.3 Program 3: PQSASA1.SAS

The PQSASA1 program processes hospital discharge abstract data and flags inpatient records if they contain outcomes of interest. Outcome indicator names have prefix of "T" (Top). Stratifier variables are constructed at the beginning of the program.

This program assumes that the input data file (consisting of inpatient discharge abstract data) conforms to specific variable names, attributes, and coding conventions. See Table 4 on page 10 for variable names and attributes for the input data file.

Partial definitions of the Prevention Quality Indicators are given in Table 1 on page 5. This table is presented to assist those individuals who desire to examine the software source code statements. Complete definitions of the indicators are given in *Prevention Quality Indicators Technical Specifications*.

Input:

1. User supplied SAS inpatient data set consisting of administrative hospital discharge abstract data. This data set is a discharge level file with an array of diagnosis and procedure codes, among other data elements.

2. SAS format library (FMTLIB) created from the PQFMTS program.

Output:

1. SAS dataset containing inpatient records with input variables, stratifiers, and flag indicators (TAPQxx) for the outcomes of interest that will later form the numerators for the PQI rates.

2. PROC MEANS (with N, NMISS, MIN, MAX, MEAN, and SUM) of all of the numeric variables in the output data file and a PROC CONTENTS of the output data file are output to hardcopy.

3. A PROC PRINT of the first 24 observations in the output data file is output to hardcopy.

Changes:

1. The "FILENAME CONTROL" statement at the beginning of the program must be changed to reflect the location of the CONTROL_PQI.SAS code.

2. In CONTROL_PQI.SAS, specify the locations (paths) of the input and output files in the "LIBNAME IN1" and "LIBNAME OUT1" statements.

3. In CONTROL_PQI.SAS, specify the names of the input and output files in the "INFILE1" and "OUTFILE1" statements.

4. The global MALEVL parameter in CONTROL_PQI.SAS by default instructs all programs to define geographic areas by county for all areas in the input datasets. Refer to Page 8 for a list of values for the MALEVL parameter.

5. *The PQI module assumes that your starting SAS data set contains 30 diagnoses, and 30 procedures. If these numbers are different in your input inpatient data, then modify the NDX and NPR parameters in CONTROL_PQI.SAS.*

6. *If your variables are named differently from those used in the PQI software (see Table 4 on page 10), you MUST either rename those variables prior to running this program or modify PQSASA1 to include the renaming of your variables to match those expected by the software. Similarly, if your variables are not the same type, i.e., character or numeric, as*

those expected by the software (see Table 4), you MUST modify your variables prior to running this program.

8.4 Program 4: PQSASA2.SAS

The PQSASA2 program calculates the observed or raw rates for the area-level Prevention Quality Indicators, using the data derived in the previous step (PQSASA1). These observed rates are stratified by combinations of area, sex, age, and race categories. The program first totals the indicator flags created by the PQSASA1 program, and then for each of the desired stratifiers divides these totals by the pertinent residential population. The population denominators are stored in variables with names that have a prefix of "P" (Pop). The Observed rates are stored in variables that have a prefix of "O".

Input:

1. The SAS dataset that was created in the previous step (PQSASA1). This is a discharge-level file that contains inpatient records with input variables, stratifiers, and the 16 new flag indicators for the PQI outcomes of interest.

A text file with Census area residential populations, stratified by area, age, sex, and ethnicity categories. One such file is currently provided along with the PQI module software. The file is POP95TO06.TXT. The user should select the year (POPYEAR) that best matches the user's discharge data file.

*Users do **not** need to convert the ASCII text file to a SAS dataset for use with the software.*

2. SAS format library (FMTLIB) created from the PQFMTS program.

Output:

1. SAS dataset with summary records that contain observed rates (OAPQxx variables where xx refers to the indicator number), the counts of outcomes that formed the numerators of the rates (TAPQxx variables), and the residential population totals that formed the denominators of the observed rates (PAPQxx variables). The output file has records summarized to the various combinations of stratifiers specified in the TYPELEVL parameter that is described in item #5 in the "Changes" section below.

2. An optional ASCII (comma delimited) text file that the user can then import into a spreadsheet. This text file contains the same information as the SAS output dataset. The text file will contain the OAPQxx observed rates, the TAPQxx counts of outcomes that formed the numerators of the rates, and the PAPQxx residential population totals. The different records/rows in the text file will correspond to the different TYPELEVL stratification combinations.

3. A PROC MEANS (with N, NMISS, MIN, MAX, MEAN, and SUM) of the area-level summary records that shows statistics for the OAPQxx observed rates, the TAPQxx counts of outcomes that formed the numerators of the rates, and the PAPQxx residential populations totals. These means will only be generated if the user included a value of 8 for the TYPELEVL parameter discussed in item 5 in the "Changes" section below.

4. A PROC CONTENTS of the output SAS summary dataset is output to hardcopy.

5. A PROC PRINT of the output summary dataset is provided at the end of the hardcopy printout. This printout may be quite large depending on the number and the types of stratifications that the user requests with the TYPELEVL parameter discussed in item 5 in the "Changes" section below. If the user does not wish to generate this printout, then the global "PRINT" parameter in CONTROL_PQI.SAS code should be set to zero.

Changes:

1. The “FILENAME CONTROL” statement at the beginning of the program must be changed to reflect the location of the CONTROL_PQI.SAS code.

2. In CONTROL_PQI.SAS, specify the locations (paths) of the input and output SAS files in the “LIBNAME IN2” and “LIBNAME OUT2” statements.

3. In CONTROL_PQI.SAS, specify the names of the input and output files in the “INFILE2” and “OUTFILA2” statements. The name that you specify with INFILE2 should be for the file that you created using the previous PQSASA1 program.

4. In CONTROL_PQI.SAS, specify the path (including the file name) for the ASCII population text file in the "FILENAME POPFILE" statement. The file POP95TO06.TXT is provided along with the PQI module. The user should set POPYEAR to the year that best matches the user’s discharge data file.

5. In CONTROL_PQI.SAS, specify the levels of aggregation for the observed rates. This is done by specifying numbers between 0 and 15 in the “%LET TYPELEVL =” statement. Each number from 0 to 15 corresponds to a different level or combination of levels as show in Table 6. The default values of 0, 8 will provide an overall total and area-level totals.

   WARNING: *TYPELEVL must include the values 0 (overall) and 8 (area) if you wish to subsequently run program PQSASA3 for risk-adjusted and smoothed rates.*

   If TYPELEVL includes the value 0, then the first observation in the output summary data file will contain the overall totals and observed rates for the entire database created with the PQSASA1 program.

Table 6. PQSASA2 aggregation possibilities

TYPELEVL	Stratification						
0	**Overall**						
1							Race
2					Sex		
3					Sex	*	Race
4			Age				
5			Age	*			Race
6			Age	*	Sex		
7			Age	*	Sex	*	Race
8	**Area**						
9	Area	*					Race
10	Area	*			Sex		
11	Area	*			Sex	*	Race
12	Area	*	Age				
13	Area	*	Age	*			Race
14	Area	*	Age	*	Sex		
15	Area	*	Age	*	Sex	*	Race

0 and 8 TYPELEVL (overall and area) are default values.

The global MALEVL parameter in CONTROL_PQI.SAS by default instructs all programs to define geographic areas by county for all areas in the input datasets. Refer to Page 8 for a list of values for the MALEVL parameter.

For example, if you are using a state inpatient hospital database, you might specify the following TYPELEVL values:

0 – provides overall rates for your entire state
4 – provides overall rates for your state, broken down by age groups
8 – provides rates for areas within your state
12 – provides rates for age groups within these areas

6. In CONTROL_PQI.SAS, if you wish to create an ASCII (comma delimited) text file that you can import into a spreadsheet, then set the TEXTA2 parameter to 1. Otherwise, the default value of zero for TEXTA2 will skip around the code that would generate this file.

7. In CONTROL_PQI.SAS, if you wish to create an ASCII (comma delimited) text file that you can import into a spreadsheet, then specify the path (including the file name) for this ASCII text file in the "FILENAME PQTEXTA2" statement.

8.5 Program 5: PQSASA3.SAS

The PQSASA3 program calculates age and sex risk-adjusted rates for each PQI (overall rates and rates by area), and then calculates smoothed rates.

Input:

1. SAS format library (FMTLIB) created from the PQFMTS program.

2. The discharge-level file that was created with the PQSASA1 program.

3. The SAS dataset with summary records that was created with the PQSASA2 program.

4. A text file with Census area residential populations, stratified by area, sex, and discrete age categories. One such file is currently provided along with the PQI module software. The file is POP95TO06.TXT. The user should select the year (i.e., POPYEAR) that best matches the user's discharge data file. The text file does not need to be converted to a SAS file for use with the software.

5. A text file (COVPQA30.TXT) containing regression coefficients from a regression that was run on a reference SID dataset. These coefficients will be used in the risk adjustment process. This file is provided to you as part of the PQI module. The text file does not need to be converted to a SAS file for use with the software.

6. A text file (MSXPQA30.TXT) containing three arrays for use in the smoothing process. The arrays contain noise estimates, signal estimates, and mean area rates for each PQI. This file is provided to you as part of the PQI module. The text file does not need to be converted to a SAS file for use with the software.

Output:

1. A SAS dataset with an overall summary record and with area-level summary records that contain the three types of indicator rates, along with the components of the initial raw rates. Specifically, the file contains the observed rates (OAPQxx variables), the risk-adjusted rates (RAPQxx variables), the smoothed rates (SAPQxx variables), the counts of outcomes that formed the numerators of the observed rates (TAPQxx variables), and the residential population totals that formed the denominators of the observed rates (PAPQxx variables).

2. An optional ASCII (comma delimited) text file that the user can then import into a spreadsheet. This text file contains the same information as the SAS output dataset. The text file will contain the OAPQxx observed rates and their TAPQxx and PAPQxx components, the RAPQxx risk-

adjusted rates, and the SAPQxx smoothed rates. Each record or row in the text file will correspond to a different area.

3. A PROC MEANS (with N, NMISS, MIN, MAX, MEAN, and SUM) of the area-level summary records that shows statistics for the OAPQxx observed rates and their TAPQxx and PAPQxx components, the RAPQxx risk-adjusted rates, and the SAPQxx smoothed rates. These means will only be generated if the user included a value of 8 for the TYPELEVL parameter discussed in item 5 in the "Changes" section for the previous PQSASA2 program.

4. A PROC CONTENTS of the output SAS summary dataset is output to hardcopy.

5. A PROC PRINT of the output summary dataset is provided at the end of the hardcopy printout. If the user does not wish to generate this printout, then the global "PRINT" parameter in CONTROL_PQI.SAS code should be set to zero.

Changes:

1. The "FILENAME CONTROL" statement at the beginning of the program must be changed to reflect the location of the CONTROL_PQI.SAS code.

2. In CONTROL_PQI.SAS, specify the locations (paths) of the input and output SAS files in the "LIBNAME IN3" and "LIBNAME OUT3" statements.

3. In CONTROL_PQI.SAS, specify in the "INFILEA3" statement the name of the discharge-level file that you created with the PQSASA1 program.

4. In CONTROL_PQI.SAS, specify the path (including the file name) for the ASCII population text file in the "FILENAME POPFILE" statement. The file name will be POP95TO06.TXT. This file is provided along with the software. The column used should be the one that is closest to the year associated with your particular data (specify POPYEAR). These files contain Census residential estimates by area, gender, and discrete age categories.

5. In CONTROL_PQI.SAS, specify in the "FILENAME COVARA" statement the path (including the file name) for the ASCII text file that contains regression coefficients from a regression that was run on reference SID dataset. The name of the file is COVPQA30.TXT.

6. In CONTROL_PQI.SAS, specify in the "FILENAME MSX" statement the path (including the file name) for the ASCII text file that contains the estimates to smooth the indicator rates. The name of the file is MSXPQA30.TXT.

7. In CONTROL_PQI.SAS, specify in the "INFILEA2" statement the name of the summary file that you created with the PQSASA2 program.

8. In CONTROL_PQI.SAS, specify in the "OUTFILA3" statement the name that you wish the output file to have.

9. In CONTROL_PQI.SAS, if you wish to create an ASCII (comma delimited) text file that you can import into a spreadsheet, then set the TEXTA3 parameter to 1. Otherwise, the default value of zero for TEXTA3 will skip around the code that would generate this file.

10. In CONTROL_PQI.SAS, if you wish to create an ASCII (comma delimited) text file that you can import into a spreadsheet, then specify the path (including the file name) for this ASCII text file in the "FILENAME PQTEXTA3" statement.

8.6 Program 6: PQSASC2.SAS

The PQSASC2 program calculates condition-specific rates for the four diabetes area-level Prevention Quality Indicators (PQIs 1, 3, 14, and 16), using the data derived in the previous step (PQSASA1). These condition-specific rates are stratified by state and age categories. The program first totals the indicator flags created by the PQSASA1 program, and then for each of the stratifiers divides these totals by the pertinent condition-specific population. The condition-specific denominators are stored in variables with names that have a prefix of "P" (Pop). The condition-specific rates are stored in variables that have a prefix of "C".

Input:

1. The SAS dataset that was created in the previous step (PQSASA1). This is a discharge-level file that contains inpatient records with input variables, stratifiers, and the 4 new flag indicators for the selected PQI outcomes of interest.

2. A text file with diabetes populations, stratified by state and age categories. Seven such files are currently provided along with the PQI module software. The files are QICTY**C**97, QICTY**C**98, QICTY**C**99, QICTY**C**00, QICTY**C**01, QICTY**C**02, QICTY**C**03, QICTY**C**04, QICTY**C**05, and QICTY**C**06. The user should select the file for the year that best matches the user's discharge data file.

   *Users do **not** need to convert the ASCII text file to a SAS dataset for use with the software.*

3. SAS format library (FMTLIB) created from the PQFMTS program.

Output:

1. SAS dataset with summary records that contain condition-specific rates (CAPQxx variables where xx refers to the indicator number), the counts of outcomes that formed the numerators of the rates (TAPQxx variables), and the condition-specific population totals that formed the denominators of the observed rates (PAPQxx variables). The output file has records for the overall state rate and by age.

2. An optional ASCII (comma delimited) text file that the user can then import into a spreadsheet. This text file contains the same information as the SAS output dataset. The text file will contain the CAPQxx condition-specific rates, the TAPQxx counts of outcomes that formed the numerators of the rates, and the PAPQxx condition-specific population totals. The different records/rows in the text file will correspond to the overall state rate and by age.

3. A PROC MEANS (with N, NMISS, MIN, MAX, MEAN, and SUM) of the area-level summary records that shows statistics for the CAPQxx condition-specific rates, the TAPQxx counts of outcomes that formed the numerators of the rates, and the PAPQxx condition-specific populations totals. These means will be generated for the overall state rate and by age.

4. A PROC CONTENTS of the output SAS summary dataset is output to hardcopy.

5. A PROC PRINT of the output summary dataset is provided at the end of the hardcopy printout. If the user does not wish to generate this printout, then the global "PRINT" parameter in CONTROL_PQI.SAS code should be set to zero.

Changes:

1. The "FILENAME CONTROL" statement at the beginning of the program must be changed to reflect the location of the CONTROL_PQI.SAS code.

2. In CONTROL_PQI.SAS, specify the locations (paths) of the input and output SAS files in the “LIBNAME INC2” and “LIBNAME OUTC2” statements.

3. In CONTROL_PQI.SAS, specify the names of the input and output files in the “INFILEC2” and “OUTFILC2” statements. The name that you specify with INFILEC2 should be for the file that you created using the previous PQSASA1 program.

4. In CONTROL_PQI.SAS, specify the path (including the file name) for the ASCII population text file in the "FILENAME POPFILEC2" statement. The file name will be QICTY**C**97, QICTY**C**98, QICTY**C**99, QICTY**C**00, QICTY**C**01, QICTY**C**02, QICTY**C**03, QICTY**C**04, QICTY**C**05, or QICTY**C**06. These files are provided to you as part of the PQI module. The file used should be the one that is closest to the year associated with your particular data.

5. In CONTROL_PQI.SAS, if you wish to create an ASCII (comma delimited) text file that you can import into a spreadsheet, then set the TEXTC2 parameter to 1. Otherwise, the default value of zero for TEXTC2 will skip around the code that would generate this file.

6. In CONTROL_PQI.SAS, if you wish to create an ASCII (comma delimited) text file that you can import into a spreadsheet, then specify the path (including the file name) for this ASCII text file in the "FILENAME PQTEXTC2” statement.

9.0 Reviewing the Printed Output

This section contains tips for reviewing some of the printed output from the PQI module. These tips are oriented toward explaining the interrelationships between printout items from different programs and hopefully will help to reveal the nature and structure of the module outputs. For guidance in interpreting the results, see section "Using Different Types of QI Rates" in the document *Guide to Prevention Quality Indicators*.

9.1 PQSASA1.SAS

The initial printout from the PQSASA1 program contains **PROC MEANS** for all of the numeric variables in the output discharge-level dataset. It will contain information for the newly constructed TAPQxx flag variables that will later form the numerators for the indicator rates. For each TAPQxx flag variable:

- The SUM will contain the total number of observations in the dataset that were found to have the particular outcome of interest.

- For the majority (13 of 14) of the indicators, the MEAN, MINIMUM, and MAXIMUM will usually be the value one since the flag variables have either been set to missing (‘.’) or to a value of one. The exception to this is PQI #2 for perforated appendix that is based on a subset of the hospitalized population rather than the area residential population (see Table 1 on page 5). For this indicator, a value of zero was assigned to the TAPQxx flag if a particular observation was part of the population for the rate denominator but did not have the particular outcome of interest to be included in the rate numerator. So for example, TAPQ02 = 0 implies a patient who had an appendectomy performed, but did not have a perforated appendix.

- For PQI #2, the MEAN will contain a close approximation of the eventual overall observed indicator rate. The value will change slightly after PQSASA2 has applied additional parts of the indicator definition.

- N lists the number of observations in the dataset with non-missing values. For the 13 area-based indicators, N for TAPQxx will be the same as the SUM. For the hospital-based indicator, N will contain the denominator for the observed indicator rate.

- There may be differences in the output from PQSASA1.SAS and PQSASA2.SAS and PQSASA3.SAS programs based on missing data. If any cases are missing the FIPS codes they will be included in the output from PQSASA1.SAS but will be excluded from the subsequent analyses (the second and third programs).

The PQSASA1.SAS printout also contains a **PROC CONTENTS** that lists all of the variables in the output dataset, including the character variables that are present, and a **PROC PRINT** of the first 24 observations in the output dataset.

9.2 PQSASA2.SAS

The purpose of this printout is to provide the observed rates for all area-level indicators. The default printout shows results by area, but other levels of output (as shown in Table 6) can be specified as well. The global MALEVL parameter in CONTROL_PQI.SAS by default instructs all programs to define geographic areas by county for all areas in the input datasets. Refer to Page 8 for a list of values for the MALEVL parameter. Note that risk-adjusted, expected, and smoothed rates are shown in the output from the next program.

The printout from the PQSASA2 program contains **PROC MEANS** on the area-level summary records for all of the numeric variables in the output summary dataset. It will contain information for the newly constructed OAPQxx rates, the PAPQxx denominators, and the TAPQxx numerators.

- The _TYPE_ variable described in the first row of the MEANS table identifies the stratification level for the records in the output data set. The _TYPE_ variable corresponds to the TYPELEVL parameter values that were specified (see item 5 in the "Changes" section for the PQSASA2 program on page 19). In this case, _TYPE_ always assumes the value of 8, since only the area-level records are selected.

- The N statistic for _TYPE_ contains the number of areas in the output summary data set. A TAPQxx numerator variable with a value for N that is lower than the N value for _TYPE_ indicates that there were no outcomes of interest for some of the areas.

- The MINIMUM value for the perforated appendix TAPQxx numerator will usually be zero since values of zero were assigned for observations that were part of the population for the rate denominator but did not have the particular outcome of interest to be included in the rate numerator. For the other 12 indicators, based on the residential area population, the MINIMUM value will be one or higher.

*For the observed rates, the User should **NOT** report the MEANS displayed here, but instead should refer to the overall means in the output dataset that are also present in the subsequent PROC PRINT. The MEANS given here are means of area means; i.e., all areas are given equal weight, regardless of the populations in the different areas.*

- The SUMs for the counter variables TAPQxx and PAPQxx will contain overall file totals for these data elements. The SUMs for the observed rates have no intuitive meaning

If the "PRINT" parameter in the CONTROL_PQI.SAS program is set to one, the subsequent **PROC PRINT** output contains a complete dump of the output summary file. Listed for each record are the stratification values, the PAPQxx population denominators for the 14 indicators, the TAPQxx outcome numerators for the 14 indicators, and the 14 OAPQxx observed rates.

9.3 PQSASA3.SAS

*The **PROC PRINT** at the end of this program provides your **final output**. (This printout appears if the "PRINT" parameter in the CONTROL_PQI.SAS program is set to one.)*

The following table lists the column headers shown on the printed output and describes each. Each indicator is reported in a separate section. Values of "xx" correspond to the two-digit PQI number.

Table 7. Area Level Printed Output

Column Heading	Description
OBS	Observation 1 is the overall average for the entire dataset (TYPELVL = 0). The remaining observations are individual areas (TYPELVL = 8).
AREA	Identifier in the dataset for area.
TAPQxx	Number of cases in the numerator.
PAPQxx	Number of cases in the denominator (population at risk).
OAPQxx	Observed (raw) rate - Numerator divided by Denominator.
RPPQxx	Risk-adjusted rate - Accounts for the difference between the case-mix of the reference population and the provider's case-mix.
LPPQxx	Lower limit of confidence interval of risk-adjusted rate.
UPPQxx	Upper limit of confidence interval of risk-adjusted rate.
SPPQxx	Smoothed rate - Rates with removal of fluxuation over time.
XPPQxx	Standard error of smoothed rate.
EPPQxx	Expected rate - Rates the area would have if it performed the same as the reference population given the area's actual case-mix (e.g., age, gender, modified DRG and comorbidities)

You may wish to express the results in more understandable terms:

- Multiply the area rates by 100,000 to express them as a rate per 100,000 population (e.g., 0.000494 * 100,000 = 49.4 cases of Congestive Heart Failure (CHF) per 100,000 population).

The MEANS table is generated from the permanent area-based output file. The means are provided just for the area-level summary records.

- The N for all but the TAPQxx counter variables should contain the number of different areas (Metro Areas and/or counties) in your database.

- The MEANs, MINIMUMs, and MAXIMUMs have their normal meaning and provide comparisons among the different areas in your database. Note that the MAXIMUMs for the counter variables (the TAPQxx and PAPQxx variables) are associated with specific areas, and therefore these MAXIMUMs may not match those in the prior PQSAS2 printouts since that run typically will include a record for the entire database.

*For the (observed, risk-adjusted, or smoothed) rates, the user should **NOT** report the MEANS displayed here, but instead refer to the overall means in the output dataset that are also present in the subsequent PROC PRINT. The MEANS given here are means of area means; i.e., all areas are given equal weight, regardless of the number of discharges in the different areas.*

- The SUMs of the counter variables (the TAPQxx and PAPQxx variables) yield the overall database totals.

9.4 PQSASC2.SAS

The purpose of this printout is to provide the condition-specific rates for selected area-level indicators. The default printout shows results by state and age category. The printout from the PQSASC2 program contains **PROC MEANS** on the state-level summary records for all of the numeric variables in the output summary dataset. It will contain information for the newly constructed CAPQxx rates, the PAPQxx denominators, and the TAPQxx numerators.

- The _TYPE_ variable described in the first column of the MEANS table identifies the stratification level for the records in the output data set. The _TYPE_ value of 1 corresponds to the state level, while a value of 3 corresponds to the state*age category level.

- The N statistic contains the number of states in the output summary data set.

- The MINIMUM value will be one or higher (for PAPQxx and TAPQxx) or zero or higher (for CAPQxx).

*For the condition-specific rates, the user should **NOT** report the MEANS displayed here, but instead should refer to the overall means in the output dataset that are also present in the subsequent PROC PRINT. The MEANS given here are means of area means; i.e., all areas are given equal weight, regardless of the populations in the different areas.*

- The SUMs for the counter variables TAPQxx and PAPQxx will contain overall file totals for these data elements. The SUMs for the *condition-specific* rates have no intuitive meaning

If the “PRINT” parameter in the CONTROL_PQI.SAS program is set to one, the subsequent PROC PRINT output contains a complete dump of the output summary file. Listed for each record are the stratification values, the PAPQxx population denominators for the 4 indicators, the TAPQxx outcome numerators for the 16 indicators, and the 4 CAPQxx condition-specific rates.

10.0 Benchmark Timings

The benchmark runtimes given below are from runs made on a Windows XP workstation, with a Pentium 4 CPU 2.80 GHz processor, 512 MB of RAM, and an IDE hard drive. The dataset used contained 757,624 observations with 15 diagnosis fields and 15 procedure fields.

Step	Run time (in seconds)
Run PQFMTS to create the format library	2
Run PQSASA1	66
Run PQSASA2	16
Run PQSASA3	55
Run PQSASC2	3

11.0 User Support

We would like to receive your feedback on the AHRQ Quality Indicators.

Our Internet address for user feedback is: support@qualityindicators.ahrq.gov

We offer a listserv to keep you informed on the Quality Indicators. We encourage you to sign up for this free service. All you need is a computer, Internet access, and an E-mail address. It works just like other electronic distribution lists.

Here's how to register:

1. Send an E-mail message to: listserv@qualityindicators.ahrq.gov.
2. On the subject line, type: Subscribe. For example:

   Subscribe
3. In the body of the message type: sub Quality_Indicators-L and your full name. For example:

   sub Quality_Indicators-L John Doe
4. You will receive a message confirming that you are signed up.

If you have any questions, contact AHRQ QI Support at the e-mail noted above. If you do not receive a confirmation message call (888) 512-6090.

Appendix A: Links

The following links may be helpful to users of the AHRQ Prevention Quality Indicators.

Prevention Quality Indicators Version 3.0a Documents and Software

Available at http://www.qualityindicators.ahrq.gov/pqi_download.htm

Title	***Description***
Guide to Prevention Quality Indicators	Describes how the PQIs were developed and provides detailed evidence for each indicator.
Prevention Quality Indicators Technical Specifications	Provides detailed definitions of each PQI, including all ICD-9-CM and DRG codes that are included in or excluded from the numerator and denominator. Note that exclusions from the denominator are automatically applied to the numerator.
PQI Covariates used in Risk Adjustment	Tables for each PQI provide the stratification and coefficients used to calculate the risk-adjusted rate for each strata.
SAS® PQI Software Documentation	This software documentation provides detailed instructions on how to use the SAS ® version of the PQI software including data preparation, calculation of the PQI rates, and interpretation of output.
SPSS® PQI Software Documentation	This software documentation provides detailed instructions on how to use the SPSS® version of the PQI software including data preparation, calculation of the PQI rates, and interpretation of output.
Change Log to PQI Documents and Software	The Change Log document provides a cumulative summary of all changes to the PQI software, software documentation, and other documents made since the release of version 2.1 of the software in March 2003. Changes to indicator specifications that were not a result of new ICD-9-CM and DRG codes, are also described in the Change Log.
Fiscal year 2006 Coding Changes	This document summarizes the changes to the indicator definitions resulting from FY 2006 changes to ICD-9-CM coding and DRG changes. These changes will only affect data from FY 2006 (October 1, 2005) or later.
SAS® PQI Software	Requires the SAS® statistical program distributed by the SAS Institute, Inc. The company may be contacted directly regarding the licensing of its products: http://www.sas.com
SPSS® PQI Software	Requires the SPSS® statistical program distributed by SPSS, Inc. The company may be contacted directly regarding the licensing of its products: http://www.spss.com

AHRQ QI Windows Application

The AHRQ QI Windows Application calculates rates for all of the AHRQ Quality Indicators modules and does not require either SAS® or SPSS®. It is available at:

http://www.qualityindicators.ahrq.gov/winqi_download.htm

Additional Documents

The following documents are available within the "Documentation" section of the **AHRQ QI Downloads** Web page:

http://www.qualityindicators.ahrq.gov/downloads.htm

- *Refinement of the HCUP Quality Indicators (Technical Review), May 2001*
- *Refinement of the HCUP Quality Indicators (Summary), May 2001*
- *Measures of Patient Safety Based on Hospital Administrative Data - The Patient Safety Indicators, August 2002*
- *Measures of Patient Safety Based on Hospital Administrative Data - The Patient Safety Indicators (Summary), August 2002*

In addition, these documents may be accessed at the AHRQ QI Documentation Web page:

http://www.qualityindicators.ahrq.gov/documentation.htm

- *Guidance for Using the AHRQ Quality Indicators for Hospital-level Public Reporting or Payment, August 2004*
- *AHRQ Summary Statement on Comparative Hospital Public Reporting, December 2005*
- *Appendix A: Current Uses of AHRQ Quality Indicators and Considerations for Hospital-level*
- *Comparison of Recommended Evaluation Criteria in Five E*xisting National Frameworks

The following documents can be viewed or downloaded from the page:

http://www.qualityindicators.ahrq.gov/newsletter.htm

- *2006 Area Level Indicator Changes*
- *Considerations in Public Reporting for the AHRQ QIs*
- *June 2005 Newsletter* - Contains the article, "Using Different Types of QI Rates"

Other Tools and Information

PQI rates can be calculated using the modified Federal Information Processing Standards (FIPS) State/county code. A list of codes is available at:

http://www.census.gov/popest/geographic/codes02.pdf

AHRQ provides a free, on-line query system based on HCUP data that provides access to health statistics and information on hospital stays at the national, regional, and State level. It is available at:

http://hcup.ahrq.gov/HCUPnet.asp

The CDC National Diabetes Surveillance System provides state level estimates of diabetes prevalence by age.

http://www.cdc.gov/diabetes/statistics/index.htm