



# **Pediatric Quality Indicators Software Instruction, SAS Version 4.3 (August 2011)**

**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. 290-04-0020 (AHRQ SQI-II)**

**Prepared by:**

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

**August 2011  
(Revised 9/1/2011)**

This document is based on research conducted by the Battelle under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-04-0020). 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 may be used, in whole or in part, as the basis for the development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

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:**

Pediatric Quality Indicators Software Instruction, SAS, Version 4.3 (Prepared by Battelle, under Contract No. 290-04-0020). Rockville, MD: Agency for Healthcare Research and Quality. August 2011.

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

## Abstract

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 for Pediatric Quality Indicators (PDI). These indicators capture volume and mortality for selected procedures, potentially preventable complications and iatrogenic events for children treated in hospitals and potentially preventable hospitalizations for ambulatory care sensitive conditions. They can be used as a screening tool for problems that patients experience as a result of exposure to the healthcare system and that are likely amenable to prevention by changes at the system or provider level.

This document is the software instruction for the Pediatric Quality Indicators software Version 4.3, 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.

Detailed definitions of the PDI, with complete listings of the ICD-9-CM and DRG codes specifically included or excluded, are contained in the document, *Pediatric Quality Indicators Technical Specifications*. *Pediatric Quality Indicators Comparative Data* provides provider rates and standard deviation, area rates and standard deviation, and population rate, as appropriate, for each indicator. See [Appendix A](#) for links to documents and tools.

## Table of Contents

Abstract .....	iv
1.0 Introduction .....	1
2.0 Components of the Pediatric Quality Indicators Module .....	3
3.0 Quick Reference .....	4
3.1 Processing Steps .....	4
3.2 Pediatric Quality Indicator Module Details .....	7
4.0 Background .....	9
5.0 Data Elements and Coding Conventions .....	11
5.1 Input for the AHRQ Quality Indicators .....	11
5.2 Coding of Diagnoses, Procedures, DRGs and MDCs .....	13
5.3 Use of External Cause of Injury Codes .....	14
5.4 Missing Values .....	14
5.5 Treatment of Missing Data .....	14
6.0 User Options .....	16
7.0 Producing Pediatric Quality Indicator Rates: Processing Steps .....	17
7.1 Step 1. Identify Outcomes in Inpatient Records .....	17
7.2 Step 2. Identify Populations at Risk .....	17
7.3 Step 3. Calculate Observed (raw) Pediatric Quality Indicator rates .....	17
7.4 Step 4. Risk Adjust the Pediatric Quality Indicator Rates .....	18
7.5 Step 5. Create MSX Smoothed Rates .....	18
8.0 Program Descriptions .....	19
8.1 Program 1: Modify Statements in the CONTROL_PDI.SAS code .....	19
8.2 Program 2: Run PDFMTS .....	20
8.3 Program 3: Run PDSAS1 .....	21
8.4 Program 4: Run PDSASG2 .....	22
8.5 Program 5: Run PDSASP2 .....	23
8.6 Program 6: Run PDSASP3 .....	28
8.7 Program 7: Run PDSASA2 .....	30
8.8 Program 8: Run PDSASA3 .....	33
9.0 Reviewing the Printed Output .....	36
9.1 PDSAS1 .....	36
9.2 PDSASP2 .....	36
9.3 PDSASP3 .....	37
9.4 PDSASA2 .....	39
9.5 PDSASA3 .....	40
10.0 User Support and Information .....	41

Appendix A: Links.....	42
Appendix B: Using Different Types of QI Rates.....	45
Appendix C: PDI Input Data Dictionary .....	47
Appendix D: SAS PDI Output Data Dictionary .....	51

## List of Figures

Figure 1. Processing steps for the pediatric quality indicators - Provider .....	5
Figure 2. Processing steps for the pediatric quality indicators - Area .....	6

## List of Tables

Table 1. Pediatric quality indicator (PDI) variables .....	7
Table 2. Prefixes for the pediatric quality indicator (PDI) variables .....	8
Table 3. Pediatric quality module contents.....	8
Table 4. Treatment of missing data elements .....	15
Table 5. PDSASP2 aggregation possibilities.....	25
Table 6. PDSASA2 aggregation possibilities .....	32
Table 7. Provider level printed output .....	38
Table 8. Area level printed output .....	40

## 1.0 Introduction

This documentation describes the software for implementing the Pediatric Quality Indicator (PDI) module of the AHRQ Quality Indicators Version 4.3 and highlights features of the analytic approach of particular interest to new users. The programs documented here were developed for use in Statistical Analysis System (SAS®)<sup>1</sup>. The SAS® version of the PDI software is designed as a PC-based single-user application. It has been tested with SAS® version 9.2 on a 32-bit machine. To use the SAS version of the QI software, users must have access to the commercially-available SAS® statistical software package, including but not limited to BASE SAS, SAS/STAT and SAS/ACCESS. For more information, visit the SAS Institute website at <http://www.sas.com>.

The software consists of a series of computer programs that perform the following functions:

- Assign and calculate PDI from hospital discharge abstracts or UB (uniform billing) claims data.
- Print PDI 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.
- The reference population database used by the software was updated to HCUP's Year 2008 State Inpatient Databases (SID). This database is used to create the regression coefficients applied in risk adjustment (see [Section 7.4](#)) and the shrinkage factors used to calculate smoothed rates (see [Section 7.5](#).)
- The syntax now supports data weighting variable (DISCWT) like those used in the Nationwide Inpatient Sample. This numeric variable must be present in the input file (see [Appendix C](#)), and should be set to 1 if weighting is not desired.
- Effective March 1, 2007 the UB-04 data specification manual includes a Present on Admission indicator (<http://www.cdc.gov/nchs/data/icd9/icdguide09.pdf>). For the PDI, POA data may impact the prevalence of the outcome of interest and the risk-adjusted rates by excluding secondary diagnoses coded as complications from the identification of covariates. If the user's input file does not include POA data, an estimation procedure is used to adjust the prevalence of the outcome of interest and the covariates to account for the missing POA data. You may include POA data for all your records, some of your records, or none of them. Including records with and without POA data is also allowed; the estimation procedure is applied to those records where POA data are missing and not to those records where POA data are not missing.
- Installation of Version 4.3 SAS syntax requires additional steps for the risk adjustment procedures to operate correctly. Once completed for one AHRQ QI module (i.e. IQI, PSI or PDI), the Prediction Module step does not need to be repeated for the other modules. The steps are:

---

<sup>1</sup> 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 QI.

- Installing the Prediction Module software
  - Click “Prediction Module” link at <http://www.qualityindicators.ahrq.gov/software/SAS.aspx>.
  - Save the “AHRQSetup 0.9.2.190.msi” file to a directory on local machine.
  - Double-click “AHRQSetup 0.9.2.190.msi” – this will install the Prediction Module executable and license files in “C:\Program Files\Battelle\AHRQ” (6 files total).
  - Create a new folder “C:\AHRQQI\AHRQPrediction\” and copy all 6 installation files to that location (so they are located in both places).
- Saving the required files to run the AHRQ QI SAS software
  - Create a sub-folder under “C:\AHRQQI\” called “PDI”.
  - The software package contains another zip file called “PDI\_GEE\_Input\_Files.zip”. All of the files within this zip file need to be extracted to the sub-folder “C:\AHRQQI\PDI”.
- The syntax will compute risk-adjusted rates and observed-to-expected ratios for the pre-defined set of stratification variables (e.g., age, gender, payer, race). When stratifications other than hospital or area are selected, the RPPDxx and RAPDxx variables and their confidence intervals are observed/expected ratios to avoid confounding with risk adjustment variables.

Version 4.3 of the PDI software includes the following changes:

- The PDI module Version 4.3 contains formats that incorporate ICD-9-CM and DRG coding through September 30, 2011. A list of the coding changes made for the fiscal year, and a *Change Log* for the PDI module may be found on the AHRQ QI website ([PDI](#)).
- The software now incorporates state level estimates of diabetes prevalence by age from the CDC National Diabetes Surveillance System. PDI #15 (Diabetes Short-term Complications Admission Rate) can be calculated using the number of diabetics in the state as the denominator, stratified by age.

Detailed information about the PDI is contained in the companion document [Measures of Pediatric Indicators](#) which includes information about the development of the indicators and their definitions. The [Pediatric Quality Indicators Comparative Data](#) document provides comparative values, including average volumes, provider rates, and population rates, depending on the type of indicator. The *Pediatric Indicators Technical Specifications* document lists all ICD-9-CM and DRG codes included in and excluded from the numerator and denominator of each indicator.

See [Appendix A](#) for links to these and other documents as well as Web sites that may be of interest to PDI users.

See [Appendix B](#), “Using Different Types of QI Rates” for an explanation of observed, expected, risk adjusted and smoothed rates and for guidance in the use of these rates. This documentation includes:

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

*Critical user information is highlighted with this symbol:* 

## 2.0 Components of the Pediatric Quality Indicators Module

As shown in Table 3 on page 8, the Pediatric Quality Indicators module consists of eight SAS programs and eight ASCII text files that contain auxiliary data. These programs and text files are described in the subsequent sections of this document.

*A note on the naming conventions for the PDI module program.* Most programs use the naming format of PDSASi, PDSASAi, or PDSASPi. The first two characters “PD” of the program name indicate a Pediatric 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 other versions of the software. In the programs that are specific to either Area-Level or Provider-Level Indicators, the sixth character of the program name is either “A” or “P”. The sixth character is an “A” if it is for the production of Area-Level indicator rates that use county or Office of Management and Budget (OMB) defined metropolitan or micropolitan area (Metro Area) populations as denominators. The sixth character is a “P” if it is a program that is for the production of Provider-Level Indicator rates that use subsets of the hospital discharge population. The last character (i) of the variable name designates the number of the specific program. One additional SAS program, PDSASG2.SAS, is the stratification module.

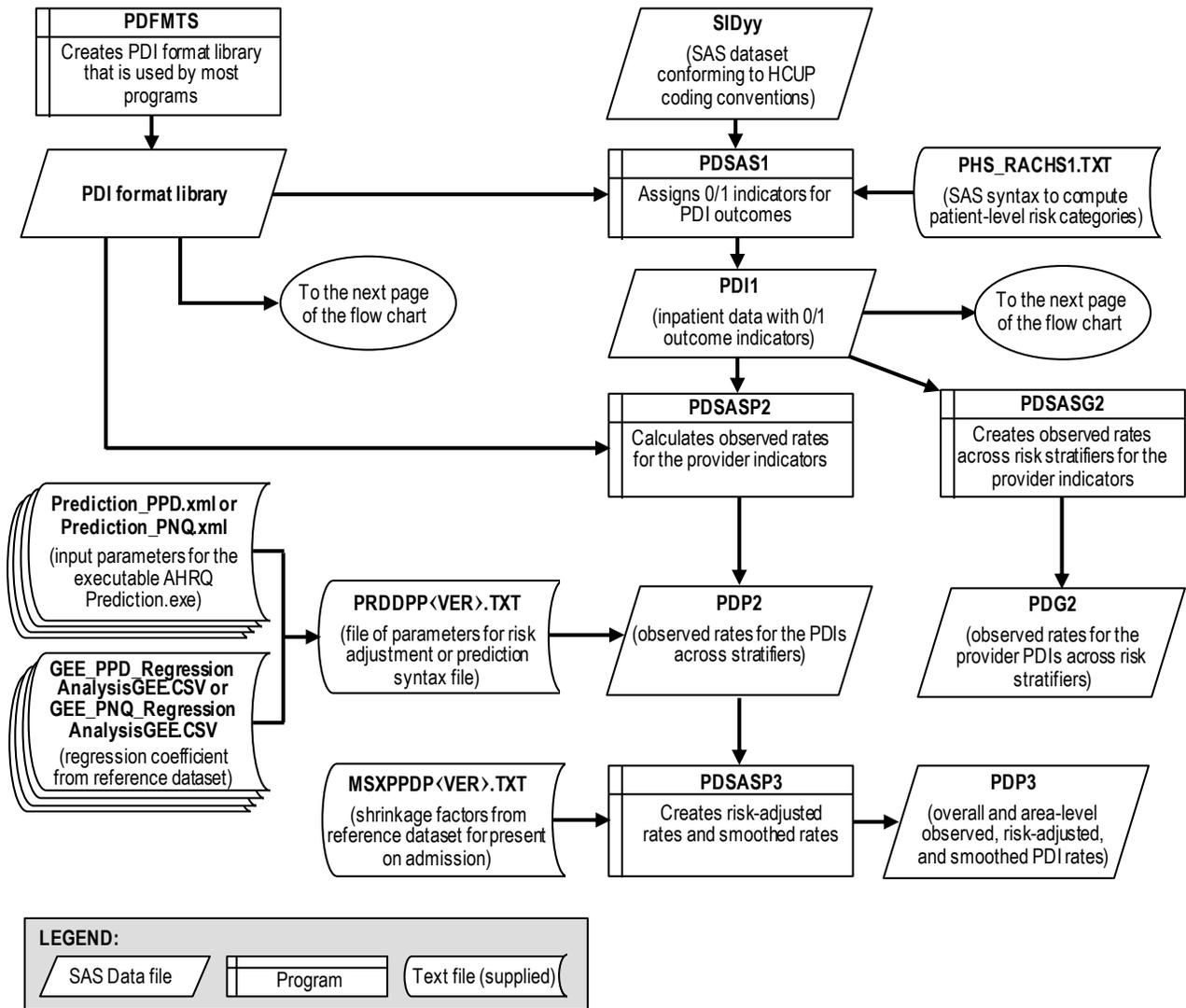
## 3.0 Quick Reference

The subsequent four pages are intended to serve as a quick reference to assist in reading this documentation and in reviewing the PDI module outputs. Processing steps are shown first (Figure 1 and Figure 2) followed by a listing of the module indicator variables (Table 1), variable prefixes (Table 2), and module contents (Table 3). One suggestion is to print these pages in duplex mode on two sheets of paper so that they can then be easily referred to as the need arises.

### 3.1 Processing Steps

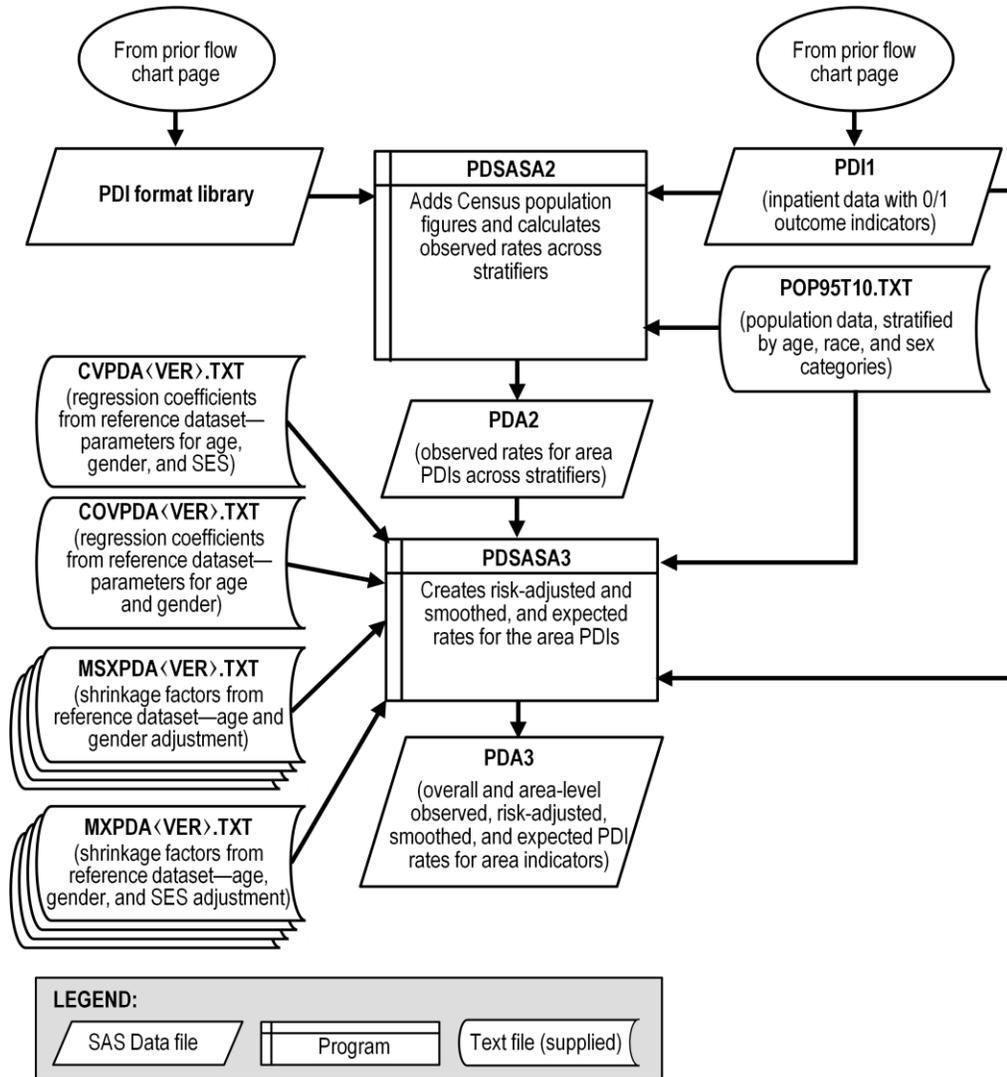
Figure 1 shows processing steps for the Pediatric Quality Indicators - Provider, while Figure 2 shows steps for the Pediatric Quality Indicators - Area.

Figure 1. Processing steps for the pediatric quality indicators - Provider



NOTE: Each of the SAS Programs in Figure 1 incorporates the CONTROL\_PDI.SAS program, which contains code the user **must** modify in order to run the PDI 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 set.

**Figure 2. Processing steps for the pediatric quality indicators - Area**



## 3.2 Pediatric Quality Indicator Module Details

The following tables list the module indicator variables (Table 1), variable prefixes (Table 2), and module contents (Table 3).

**Table 1. Pediatric quality indicator (PDI) variables**

PDI Number	Indicator
Pediatric Quality Provider Indicators	
1	Accidental Puncture or Laceration
2	Pressure Ulcer
3	Foreign Body Left During Procedure
5	Iatrogenic Pneumothorax
6	Pediatric Heart Surgery Mortality
7	Pediatric Heart Surgery Volume
8	Postoperative Hemorrhage or Hematoma
9	Postoperative Respiratory Failure
10	Postoperative Sepsis
11	Postoperative Wound Dehiscence
12	Central Line-associated Blood stream Infection
13	Transfusion Reaction
PSI 17	Birth Trauma-Injury to Neonate <sup>a</sup>
PQI 9	Low Birth Weight <sup>a</sup>
Pediatric Quality Area Indicators	
14	Asthma Admission Rate
15	Diabetes Short-term Complications Admission Rate
16	Gastroenteritis Admission Rate
17	Perforated Appendix Admission Rate
18	Urinary Tract Infection Admission Rate
90	PDI Composite
91	PDI Composite (acute only)
92	PDI Composite (chronic only)
Neonatal Quality Indicators	
1	Iatrogenic Pneumothorax in Neonates
2	Neonatal Mortality
3	Blood stream Infection in Neonates

<sup>a</sup> Indicators PSI 17 and PQI 09 are calculated by the PDI SAS module because these indicators are based on pediatric discharges.

**Table 2. Prefixes for the pediatric quality indicator (PDI) variables**

Prefix	Contents	Examples	
		Provider Indicator # 8, Postoperative Hemorrhage or Hematoma	Area Indicator # 14, Asthma
T	Inpatient numerator (top)	TPPD08	TAPD14
P	Population denominator (pop)	PPPD08	PAPD14
Q	POA flag	QPPD08	N/A
O	Observed rate	OPPD08	OAPD14
E	Expected rate	EPPD08	EAPD26
R	Risk-adjusted rate	RPPD08	RAPD26
L	Lower limit of confidence interval of risk-adjusted rate	LPPD08	N/A
U	Upper limit of confidence interval of risk-adjusted rate	UPPD08	N/A
S	Smoothed rate	SPPD08	N/A
X	Standard error of smoothed rate	XPPD08	N/A

**Table 3. Pediatric quality module contents**

SAS programs (syntax files)	ASCII text files (data)
CONTROL_PDI.SAS	POP95T11.TXT
PDFMTS.SAS	MSXPDA43.TXT
PDSAS1.SAS	MSXPDP43.TXT
PDSASA2.SAS	PHS_RACHS1.TXT
PDSASA3.SAS	gee_ppdxx_RegressionAnalysisGee.csv *
PDSASG2.SAS	Prediction_PPDxx.xml *
PDSASP2.SAS	gee_pnqxx_RegressionAnalysisGee.csv *
PDSASP3.SAS	Prediction_PNQxx.xml *
	PRDPDP43.TXT
	MXPDA43A.TXT
	MSXPDC43.TXT
	CVPDA43A.TXT
	COVPDA43.TXT

\*contained in PDI\_GEE\_Input.zip

The ASCII text files are data files provided with the Pediatric Quality 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. . Note that POP95T11.TXT is available on the website as a separate download.

## 4.0 Background

The PDI module contains 15 indicators that reflect the quality of care inside hospitals and 8 indicators that reflect the quality of care outside hospitals. Note that PQI 09 and PSI 17 are not included in the list, but these indicators are also calculated using PDI module since they are based on discharges. These indicators serve as flags for potential quality problems rather than definitive measures of quality of care. The indicators include two groups of measures based on hospital stays.

1. There are 15 **Provider-based** Pediatric Quality Indicators for volume and in-hospital mortality for selected procedures and for medical conditions and surgical procedures that have been shown to have complication/adverse event rates that vary substantially across institutions and for which evidence suggests that high complication/adverse event rates may be associated with deficiencies in the quality of care. There are also three indicators for neonates. These indicators are measured as rates—outcome of interest / adverse events divided by the population at risk for the procedure or condition. The adverse event related Provider-based Indicators include only those cases where a secondary diagnosis code flags a potentially preventable complication.
2. Four of the adverse event related indicators are for surgical discharges and seven are for either medical or surgical discharges. Four of the Provider-based Pediatric Quality Indicators (PDI 2, 8, 9, 11), incorporate information about when procedures were performed (relative to the admission date) if this information is available. Note, however, that if the day of procedure information is missing, the computed rates for these indicators will be slightly lower than if the information was available. Admission type is used by two of the PDI, postoperative hemorrhage or hematoma and postoperative respiratory failure (PDI #8 and PDI #9), to identify elective surgeries.
3. There are eight **Area-based** Pediatric Quality Indicators (five individual indicators and three composites), where geographic areas can be defined at the user's option in one of four ways:
  - County level with U.S. Census FIPS<sup>2</sup>
  - County level with modified FIPS
  - Metro Area with OMB 1999 definition
  - Metro Area with OMB 2003 definition.
4. PQI 9, Low Birth Weight, is now generated by the SAS Pediatric module: although it retains its original name. This measure uses in-hospital live births as the denominator.
5. PSI 17, Birth Trauma – Injury to Neonate, is now generated by the SAS Pediatric module although it retains its original name since it is based on pediatric discharges. This measure uses all in-hospital newborns as the denominator.

The MALEVL parameter reflects the change in OMB definitions for areas. Refer to page 13 for information on these area definitions and MALEVL parameter settings.

---

<sup>2</sup> 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 24 indicators in the PDI module are listed in Table 1 on page 7. See also *Pediatric Quality Indicators Technical Specifications* ([http://www.qualityindicators.ahrq.gov/modules/pdi\\_resources.aspx](http://www.qualityindicators.ahrq.gov/modules/pdi_resources.aspx)) for the definition of each indicator.

The data required for measuring these indicators come from hospital discharge abstracts or billing claims (administrative data) which are readily available within hospitals or from many state data organizations. The residential population data for the five Area-based Indicators are from the U.S. Census Bureau.

For information about how these indicators were identified, see *AHRQ QI Development* available at [http://www.qualityindicators.ahrq.gov/modules/pdi\\_resources.aspx](http://www.qualityindicators.ahrq.gov/modules/pdi_resources.aspx).

## 5.0 Data Elements and Coding Conventions

See [Appendix C](#) for a detailed description of the input data elements and coding conventions used by the SAS QI software.

### 5.1 Input for the AHRQ Quality Indicators

The PDI software was written to process data from discharge data abstracts that contain information about hospital stays. The PDI module was tested on data from the AHRQ 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. Additional information on HCUP is available at the website <http://www.ahrq.gov/data/hcup>.

-  *The input data files for the Pediatric Quality Indicators software must be in SAS.*
-  *It is strongly recommend that users recode data elements in their input files to be consistent with the coding expected by the software. We do not recommend changing the Pediatric Quality Indicators software. For example, recoding the SEX data element in the input file to be consistent with the coding described in [Appendix C](#) (i.e., 1 for male and 2 for female) should be done rather than attempting to modify all uses of the Sex data element in the PDI programs.*
-  *Not every value for every data element is referenced by the PDI software. For example, only two of the discharge disposition (DISP) data elements values are used in the software (DISP value “2” for short-term hospital and value “20” for died in the hospital). However, the complete set of standardized values is included, based on HCUP specifications, to assist users of the uniform HCUP data files.*
-  *Two of the indicators (PDI #8 and 9) use admission type (ATYPE) to identify elective surgeries. If admission type (ATYPE) is not available in the user’s data, the user may want to evaluate other data in the file to create a proxy to identify elective admission types.*
-  *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 PDI programs.*

The data elements listed in [Appendix C](#) are those required for the input files of **all** of the SAS QI modules. When a variable is not required for the PDI SAS program the notation “Not used by the QI 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 as well as users working to standardize their input file specifications for use with any of the AHRQ QI software modules.

 *Appendix C 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 name of the data element PSTCO emphasizes the importance of calculating the area PDI by the location of the patient residence. If the user wants to calculate the area PDI 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 PDI 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 provider 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.*

 *Modified FIPS codes assignment for area denominators is an available option. 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. Using the provider FIPS state/county code for analysis should be done with caution, and at larger levels of geographic area (i.e. Metropolitan 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 PDI area rates. They will be included in the output from the first program (PDSASI.SAS).*

 *The software will generate area PDI 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.*

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 always 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 New England County Metropolitan Area (NECMAs). OMB defines NECMAs as a county-based alternative to the city- and town-based New England MSAs and Consolidated MSAs (CMSAs). 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 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, DRGs and MDCs

 *Diagnoses and procedures must be coded using the International Classification of Diseases, 9th<sup>R</sup> revision, Clinical Modification (ICD-9-CM)*



 *Significant modifications to ICD-9-CM codes occurred in the early 1990s. PDI definitions only reflect codes valid after October 1, 1997, so they may not accurately analyze data collected before that date.*

- Diagnosis 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 PDI software 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 truncated versions at 3 or 4 digits. Similarly, if an ICD code is specified using 4 digits, the software recognizes only 4-digit versions of that code and not truncated versions 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.

- Diagnosis-Related Groups (DRGs) and Major Diagnostic Categories (MDCs) 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 and MDC effective on the discharge date. The software now refers to data elements DRG and MDC. Your data should be coded accordingly.
-  Version 4.3 of the software accounts for ICD-9-CM, DRG and MDC coding effective through FY 2011 (or through September 30, 2011).

## 5.3 Use of External Cause of Injury Codes

External cause of injury codes (E-codes ) are used to classify environmental events, circumstances, and conditions as the cause of injury, poisoning, or other adverse events. The use of E-codes is not always required by a state uniform billing committee or a state data organization. Users should be knowledgeable of the E-code requirements and practices of hospitals represented in the input data file. Several of the indicators use E-codes in their numerator or denominator definitions. See the documents *Measures of Pediatric Health Care Quality based on Hospital Administrative Data* and *Pediatric Quality Indicators Technical Specifications* ([http://www.qualityindicators.ahrq.gov/Modules/pdi\\_resources.aspx](http://www.qualityindicators.ahrq.gov/Modules/pdi_resources.aspx)) for additional information.

The data elements in the AHRQ QI are based on the coding specifications used in the State Inpatient Databases (SID) in the Healthcare Cost and Utilization Project (HCUP). The SID coding specifications are similar to the Uniform Bill (UB-04) but not identical. For data elements used in the AHRQ QI, crosswalks between the SID and UB-04 coding specifications are included in the SID documentation available at <http://hcup-us.ahrq.gov/db/state/siddbdocumentation.jsp>.

## 5.4 Missing Values

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

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 PDI programs do not distinguish among the different types of missing codes. Therefore, all types of missing values in the input data to the PDI module must be represented by a single value – missing data (.) for numeric variables and blank ( ‘ ’ ) for alphanumeric (or character) variables.

## 5.5 Treatment of Missing Data

The software is designed to handle missing data in a particular manner; 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 are 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. **Error! Reference source not found.** lists the impact of missing data for each data element.

**Table 4. Treatment of missing data elements**

Variable	Label	D	N	E	S	Treatment of Missing Data
AGE	Age In Years At Admission	X	X	X	X	Case excluded from all analysis (e.g., dropped from the denominator of the discharge based indicators and from numerator of all population based measures).
AGEDAY	Age in days (coded only when the age in years is less than 1)					The value is treated as if it was less than 29 days.
ATYPE	Admission Type	X				Case excluded from denominator where used in specification.
DISP	Disposition Status		X	X		Case excluded from denominator where used in specification (i.e., PDI #06 and NQI #02).
DQTR	Discharge Quarter	X	X	X	X	Case excluded from all analysis (e.g., dropped from the denominator of the discharge based indicators and from numerator of all population based measures).
DX1	Principal Diagnosis Code	X	X	X	X	Record excluded from analysis
LOS	Length of Stay	X		X		Case excluded from denominator where used in specification.
SEX	Patient Gender	X	X		X	Case excluded from all analysis (e.g., dropped from the denominator of the discharge based indicators and from numerator of all population based measures).
PAY1	Payer				X	Classified as "Other."
PSTCO	Location of Patient Residence or Hospital Location Modified FIPS State/County Code	X				Dropped from denominator in rate calculations (stratified by county or Metro Area), case will appear in calculation of the overall rate.
RACE	Race				X	Classified As "Other."
YEAR	Discharge Year	X	X	X	X	Case excluded from all analysis (e.g., dropped from the denominator of the discharge based indicators and from numerator of all population based measures).

Note. D = Denominator; N = Numerator; E = Exclusion; S = Stratification.

## 6.0 User Options

The PDI software reflects the development and implementation of the software with discharge data available to AHRQ through the HCUP program. 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, users are encouraged to consider how AHRQ may modify the PDI software to better serve their measurement needs and 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 PDI based on original definitions is lost. Users are encouraged to identify ways to improve the PDI methodology and to share their suggestions with AHRQ for future updates.*

AHRQ provides public access to HCUPnet - an interactive tool for identifying, tracking, analyzing, and comparing statistics on hospital care. HCUPnet queries generate statistics in a table format using data from the Nationwide Inpatient Sample (NIS) and State Inpatient Databases (SID) for those States that have agreed to participate. HCUPnet includes statistics based on the AHRQ Quality Indicators (QIs) which have been applied to the HCUP NIS. If users change definitions of the PDI, it will not be possible to compare users' results to the national estimates in HCUPnet. HCUPnet may be accessed through the HCUP web site at <http://hcup.ahrq.gov/HCUPnet.asp>.

## 7.0 Producing Pediatric Quality Indicator Rates: Processing Steps

Each Pediatric Quality Indicator (PDI) expressed as a rate, is simply defined as:

Outcome of interest / Population at risk

[or]

Numerator / Denominator

Conceptually, three steps are necessary to produce the PDI 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 “T” for “top”) for each of the AHRQ PDI measures.

Records are marked by setting a series of flag variables, each of which corresponds to the numerator for a particular PDI. For example, if the inpatient record meets the conditions for inclusion in the outcome for PDI-#01, then the outcome indicator for PDI-#01 is set to one.

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 PDI. 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

For the Provider PDI, the populations at risk (the denominators for calculating the PDI rates, or “pop”) are derived from the hospital discharge records.

For the Area PDI, the populations at risk (the denominators for calculating the PDI rates) are derived from U.S. Census population figures.

### 7.3 Step 3. Calculate Observed (raw) Pediatric Quality Indicator rates

Using the output data from Step 1 and either hospital discharge or Census population data from Step 2, the PDI rates are calculated for user-specified combinations of stratifiers. These observed rates are stratified by risk group categories that are specific to each indicator for Provider PDI, and areas (Metro areas/counties), age groups, race/ethnicity categories, and sex for the Area PDI.



*The programs calculate observed PDI rates regardless of the number of cases available. However, PDI 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 for user-specified categories, alone or in combination. For example, Provider PDI observed rates could be output at the hospital level, at the payer level, or at the hospital level by payer.

## 7.4 Step 4. Risk Adjust the Pediatric Quality Indicator Rates

Regression coefficients from a reference population database (reflecting a large proportion of the U.S. population) are applied to the observed rates in the risk-adjustment process. These reference population file regression coefficients are provided as part of the PDI module. The risk-adjusted rates will then reflect the age, sex, severity and comorbidity distribution of data in the reference population file rather than the distribution for the hospitals or areas present in the users' data. This approach will allow risk-adjusted rates produced by various users to be compared directly to one another. The regression coefficients were derived from HCUP's State Inpatient Databases (SID). The code to generate these reference population means and coefficients are not part of the PDI module.

 *The programs calculate risk-adjusted PDI rates regardless of the number of cases available (a minimum of 3 cases). However, PDI rates based on only a few cases should be interpreted with caution.*

## 7.5 Step 5. Create MSX Smoothed Rates

Shrinkage factors are applied to the risk-adjusted rates for each PDI in a process called multivariate signal extraction (MSX). These shrinkage factors were calculated from the HCUP SID. For each PDI, the shrinkage estimate reflects a 'reliability adjustment' unique to each indicator. The less reliable the PDI over time and across hospitals or areas, the more the estimate 'shrinks' the PDI 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 PDI software and do not need to be calculated by users.

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

## 8.0 Program Descriptions

This section describes the eight SAS programs that assign, calculate, and print the Pediatric 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 PDI module programs is shown in the flowcharts in Figure 1 and Figure 2 that begin on page 5.

If you desire to create and examine observed PDI rates, then the CONTROL, PDFMTS, and the PDSAS1 programs will need to be used, followed by PDSASP2 (for Provider observed rates) and/or PDSASA2 (for Area observed rates across stratifiers). The Pediatric module offers two alternatives for analyzing provider level rates, stratification and risk adjustment. If you wish to create stratified Provider PDI, you will need to run PDSASG2. If you wish to create risk-adjusted and smoothed PDI rates, then you will also need to run the PDSASP3 and/or the PDSASA3.

### 8.1 Program 1: Modify Statements in the CONTROL\_PDI.SAS code

The CONTROL\_PDI.SAS code contains all of the SAS statements that you need to modify in order to run the remaining PDI module programs (with one exception noted below). This location is where you specify, for example, the path names or locations of all your input and output data sets. There is also a TITLE line that the user can fill in. The items that need to be changed in order to run the PDSAS1, PDSASP2, PDSAP3, PDSASG2, PDSASA2. And PDSASA3 programs are described below under each program.

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

 *All of the changes needed for the CONTROL\_PDI.SAS code are listed in the following descriptions for programs 2 – 8. Each program description contains a section labeled “Changes” that enumerates the modifications for CONTROL\_PDI.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 PDI 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\_PDI.SAS file. Pathnames should not exceed 255 characters in length.

The PHS\_RACHS1.TXT text file includes the SAS syntax for implementing the Risk Adjustment for Congenital Heart Surgery, Version 1 (RACHS1) methodology. The SAS syntax is included in a separate text file only for ease of maintenance and updating. It is required syntax

and the user must designate the appropriate directory containing this text file in CONTROL\_PDI.SAS

CONTROL\_PDI.SAS has a collection of parameters that allow you to read in data from different pathnames and write out data to different locations or pathnames. For example, “LIBNAME1 IN1” points to the location of your input data set for program PDSAS1 and “LIBNAME OUT1” points to the location of the output data set created by the PDSAS1 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\_PDI.SAS changing “C:\pathname” to the single directory location that you wish to use.

There is a global PRINT parameter in CONTROL\_PDI.SAS that allows you to print all of the output summary records at the end of the PDSASP2, PDSASP3, PDSASG2, PDSASA2, and PDSASA3 programs. These printouts may be quite large if you have a large database with a large number of hospitals and if you choose the results to be finely stratified. If you wish to turn this feature off, set PRINT to have a value of zero in CONTROL\_IQI.SAS; to turn this feature on set PRINT to have a value of 1.

 *The global MALEVL parameter in CONTROL.PDI.SAS by default instructs all programs to define geographic areas by county for all areas in the input datasets. Refer to Page 13 for a list of values for the MALEVL parameter.*

## 8.2 Program 2: Run PDFMTS

The PDFMTS 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 (i.e., age, sex, race/ethnicity and payer). These are followed by formats that are used for the identification of outcomes of interest in the PDSAS1.SAS program, and by formats that map FIPS county codes to Metro Areas in the PDSAS1, PDSASA2 and PDSASA3 programs.

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 it points to your location of the CONTROL\_PDI.SAS code.
2. In CONTROL\_PDI.SAS, specify the location (path) of the format library in the “LIBNAME LIBRARY” statement.

## 8.3 Program 3: Run PDSAS1

The PDSAS1 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 15 for variable names and attributes for the input data file.

Partial definitions of the Pediatric Quality Indicators are given in Table 1 on page 7. 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 the document *Pediatric Quality Indicators Technical Specifications* ([http://www.qualityindicators.ahrq.gov/modules/pdi\\_resources.aspx](http://www.qualityindicators.ahrq.gov/modules/pdi_resources.aspx) ).

### 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 PDFMT program.

### Output:

1. SAS dataset containing inpatient records with input variables, stratifiers, comorbidity variables and flag indicators (TPPDxx and TAPDxx) for the outcomes of interest that will later form the numerators for the PDI rates. It also includes variables identifying present on admission (POA) and conditions for provider based outcomes of interest (QPPDxx).
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\_PDI.SAS code.
2. In CONTROL\_PDI.SAS, specify the locations (paths) of the input and output files in the “LIBNAME IN0” and “LIBNAME OUT1” statements.
3. In CONTROL\_PDI.SAS, specify the names of the input and output files in the “INFILE0” and “OUTFILE1” statements.

 *The PDI 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\_PDI.SAS.*

 *If your variables are named differently from those used in the PDI software (see **Error! Reference source not found.** on page 15), you **MUST** either rename those variables prior to running this program or modify PDSAS1 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 **Error! Reference source not found.**), you **MUST** modify your variables prior to running this program.*

## 8.4 Program 4: Run PDSASG2

The PDSASG2 program calculates the observed or raw rates for the Provider-level Pediatric Quality Indicators, using the data derived in a previous step (PDSAS1). These observed rates are stratified by risk group categories that are specific to each indicator. The program first sums the indicator flags created by the PDSAS1 program (numerator) and then divides this sum by the count of the hospital discharges (denominator) for each of the risk group stratifiers. The stratifiers are stored in variables with names that have a prefix of “G” (Group). The numerators are stored in variables with names that have a prefix of “T” (Top). 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 PDSAS1. This dataset is a discharge-level file that contains inpatient records with input variables, stratifiers, and the 13 new flag indicators for the PDI Hospital outcomes of interest. Only 6 of these indicators have risk group stratifiers and are included as input into the G2 module (#1, #2, #8, #10, #11 and #12).
2. SAS format library (FMTLIB) created from the PDFMT program.

### Output:

1. SAS dataset with summary records that contain observed rates (OPPDxx variables where xx refers to the indicator number), the counts of outcomes that formed the numerators of the rates (TPPDxx variables), and the hospital discharge totals that formed the denominators of the observed rates (PPPDxx variables). It also includes variables indentifying present on admission (POA) and conditions for provider based outcomes of interest (QPQIxx).The output file has records summarized to the indicator-specific pre-defined risk group stratifiers.
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 OPPDxx observed rates, the TPPDxx counts of outcomes that formed the numerators of the rates and the PPPDxx counts of outcomes

that formed the denominators of the rates. The different records/rows in the text file will correspond to the indicator-specific pre-defined risk group stratifiers

3. A PROC MEANS (with N, NMISS, MIN, MAX, MEAN and SUM) of the hospital-level summary records that shows statistics for the OPPDxx observed rates, the TPPDxx counts of outcomes that formed the numerators of the rates, the QPPDxx POA flags, and the PPPDxx counts of outcomes that formed the denominators of the rates. These means are generated by the indicator-specific pre-defined risk group stratifiers.
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\_PDI.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\_PDI.SAS code.
2. In CONTROL\_PDI.SAS, specify the locations (paths) of the input and output SAS files in the “LIBNAME IN1” and “LIBNAME OUTG2” statements.
3. In CONTROL\_PDI.SAS, specify the names of the input and output files in the “INFILE1” and “OUTFILG2” statements. The name that you specify with INFILE1 should be for the file that you created using the PDSAS1 program.
4. In CONTROL\_PDI.SAS, if you wish to create an ASCII (comma delimited) text file that you can import into a spreadsheet, then set the TEXTG2 parameter to 1. Otherwise, the default value of zero for TEXTG2 will skip around the code that would generate this file.
5. In CONTROL\_PDI.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 PDTEXTG2” statement.

## 8.5 Program 5: Run PDSASP2

The PDSASP2 program calculates the observed or raw rates for the Provider-level Pediatric Quality Indicators, using the data derived in a previous step (PDSAS1). These observed rates are stratified by combinations of hospitals, sex, age, race and payer categories. The program first totals the indicator flags created by the PDSAS1 program and then for each of the desired stratifiers divides these totals by the hospital discharges in the universe for the indicator. 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 PDSAS1. This is a discharge-level file that contains inpatient records with input variables, stratifiers, and the 13 new flag indicators for the PDI Hospital outcomes of interest.
2. SAS format library (FMTLIB) created from the PDFMT program.

Output:

1. SAS dataset with summary records that contain observed rates (OPPDxx variables where xx refers to the indicator number), the counts of outcomes that formed the numerators of the rates (TPPDxx variables), and the hospital discharge totals that formed the denominators of the observed rates (PPPDxx variables). The output file has records summarized to the various combinations of stratifiers specified in the TYPELVLP parameter that is described in item 4 in “Changes” 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 OPPDxx observed rates, the TPPDxx counts of outcomes that formed the numerators of the rates, and the PPPDxx counts of outcomes that formed the denominators of the rates. The different records/rows in the text file will correspond to the different TYPELVLP stratification combinations
3. A PROC MEANS (with N, NMISS, MIN, MAX, MEAN, and SUM) of the hospital-level summary records that shows statistics for the OPPDxx observed rates, the TPPDxx counts of outcomes that formed the numerators of the rates, and the PPPDxx counts of outcomes that formed the denominators of the rates. These means will only be generated if the user included a value of 64 for the TYPELVLP parameter discussed in item 4 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 TYPELVLP parameter discussed in item 4 in the “Changes” section below. If the user does not wish to generate this printout, then the global “PRINT” parameter in CONTROL\_PDI.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\_PDI.SAS code.
2. In CONTROL\_PDI.SAS, specify the locations (paths) of the input and output SAS files in the “LIBNAME IN1” and “LIBNAME OUTP2” statements.

3. In CONTROL\_PDI.SAS, specify the names of the input and output files in the “INFILE1” and “OUTFILP2” statements. The name that you specify with INFILE1 should be for the file that you created using the PDSAS1 program.
4. In CONTROL\_PDI.SAS, specify the levels of aggregation for the observed rates. This is done by specifying numbers between 0 and 127 in the “%LET TYPELVLP =” statement. Each number from 0 to 31 corresponds to a different level or combination of levels as show in Table 5. The default values of 0 and 64 will provide an overall total and provider-level totals.



*If TYPELVLP includes the value 0, then the first observation in the output summary datafile will contain the overall totals and observed rates for the entire database created with the PDSAS1 program.*

**Table 5. PDSASP2 aggregation possibilities**

TYPELVLP	Stratification
0	overall
1	race
2	payer
3	payer * race
4	sex
5	sex * race
6	sex * payer
7	sex * payer * race
8	age
9	age * race
10	age * payer
11	age * payer * race
12	age * sex
13	age * sex * race
14	age * sex * payer
15	age * sex * payer * race
16	ageday
17	ageday * race
18	ageday * payer
19	ageday * payer * race
20	ageday * sex
21	ageday * sex * race
22	ageday * sex * payer
23	ageday * sex * payer * race
24	ageday * age
25	ageday * age * race
26	ageday * age * payer
27	ageday * age * payer * race
28	ageday * age * sex
29	ageday * age * sex * race
30	ageday * age * sex * payer

TYPSELVLP	Stratification
31	ageday * age * sex * payer * race
32	bwht
33	bwht * race
34	bwht * payer
35	bwht * payer * race
36	bwht * sex
37	bwht * sex * race
38	bwht * sex * payer
39	bwht * sex * payer * race
40	bwht * age
41	bwht * age * race
42	bwht * age * payer
43	bwht * age * payer * race
44	bwht * age * sex
45	bwht * age * sex * race
46	bwht * age * sex * payer
47	bwht * age * sex * payer * race
48	bwht * ageday
49	bwht * ageday * race
50	bwht * ageday * payer
51	bwht * ageday * payer * race
52	bwht * ageday * sex
53	bwht * ageday * sex * race
54	bwht * ageday * sex * payer
55	bwht * ageday * sex * payer * race
56	bwht * ageday * age
57	bwht * ageday * age * race
58	bwht * ageday * age * payer
59	bwht * ageday * age * payer * race
60	bwht * ageday * age * sex
61	bwht * ageday * age * sex * race
62	bwht * ageday * age * sex * payer
63	bwht * ageday * age * sex * payer * race
64	provider
65	provider * race
66	provider * payer
67	provider * payer * race
68	provider * sex
69	provider * sex * race
70	provider * sex * payer
71	provider * sex * payer * race
72	provider * age
73	provider * age * race
74	provider * age * payer
75	provider * age * payer * race
76	provider * age * sex
77	provider * age * sex * race
78	provider * age * sex * payer

TYP	LEVEL	Stratification
79	provider	* age * sex * payer * race
80	provider	* ageday
81	provider	* ageday * race
82	provider	* ageday * payer
83	provider	* ageday * payer * race
84	provider	* ageday * sex
85	provider	* ageday * sex * race
86	provider	* ageday * sex * payer
87	provider	* ageday * sex * payer * race
88	provider	* ageday * age
89	provider	* ageday * age * race
90	provider	* ageday * age * payer
91	provider	* ageday * age * payer * race
92	provider	* ageday * age * sex
93	provider	* ageday * age * sex * race
94	provider	* ageday * age * sex * payer
95	provider	* ageday * age * sex * payer * race
96	provider	* bwht
97	provider	* bwht * race
98	provider	* bwht * payer
99	provider	* bwht * payer * race
100	provider	* bwht * sex
101	provider	* bwht * sex * race
102	provider	* bwht * sex * payer
103	provider	* bwht * sex * payer * race
104	provider	* bwht * age
105	provider	* bwht * age * race
106	provider	* bwht * age * payer
107	provider	* bwht * age * payer * race
108	provider	* bwht * age * sex
109	provider	* bwht * age * sex * race
110	provider	* bwht * age * sex * payer
111	provider	* bwht * age * sex * payer * race
112	provider	* bwht * ageday
113	provider	* bwht * ageday * race
114	provider	* bwht * ageday * payer
115	provider	* bwht * ageday * payer * race
116	provider	* bwht * ageday * sex
117	provider	* bwht * ageday * sex * race
118	provider	* bwht * ageday * sex * payer
119	provider	* bwht * ageday * sex * payer * race
120	provider	* bwht * ageday * age
121	provider	* bwht * ageday * age * race
122	provider	* bwht * ageday * age * payer
123	provider	* bwht * ageday * age * payer * race

TYPELVLP	Stratification
124	provider * bwht * ageday * age * sex
125	provider * bwht * ageday * age * sex * race
126	provider * bwht * ageday * age * sex * payer
127	provider * bwht * ageday * age * sex * payer * race

0 and 64 TYPELVLP (overall and provider) are default values for the level of computation in the software.

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

- 0 – provides overall rates for your entire state
- 8 – provides overall rates for your state, broken down by age groups
- 64 – provides rates for providers within your state
- 72 – provides rates for age groups within these providers

5. In CONTROL\_PDI.SAS, if you wish to create an ASCII (comma delimited) text file that you can import into a spreadsheet, then set the TEXTP2 parameter to 1. Otherwise, the default value of zero for TEXTP2 will skip around the code that would generate this file.
6. In CONTROL\_PDI.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 PDTEXTP2” statement.

## 8.6 Program 6: Run PDSASP3

The PDSASP3 program calculates risk-adjusted rates by patient birthweight, age, modified-DRG and CCS comorbidity for each PDI and then calculates smoothed and expected rates.

 *The programs calculate risk-adjusted PDI rates regardless of the number of cases available. However, PDI rates based on only a few cases should be interpreted with caution.*

The output file has records summarized to the various combinations of stratifiers that were specified in the TYPELVLP parameter of the preceding PDSASP2 program. See Table 5 on page 25.

### Input:

1. SAS format library (FMTLIB) created from the PDFMT program.
2. The discharge-level file that was created with the PDSAS1 program.
3. The SAS dataset with summary records that was created with the PDSASP2 program.
4. A text file containing regression coefficients from a regression that was run on a reference SID dataset. These coefficients will be used in the risk-adjustment process. Two files are provided as part of the PDI module.

Gee\_ppdxx\_RegressionAnalysisGee.csv contains the regression coefficients.  
Prediction\_PPDxx.xml contains the input parameters for the executable AHRQ.exe.

 *Users must create a directory “C:\AHRQOI\ AHRQPrediction” and copy the executable AHRQ.exe to that directory (along with 5 other files). Users must also create a directory “C:\AHRQOI\PDI” and copy the zip file PDI\_GEE\_Input\_Files.zip to that directory. Extract all of the files from PDI\_GEE\_Input\_Files.zip into the directory “C:\AHRQOI\PDI”. (See version 4.3 introduction on page 1)*

5. A text file containing two arrays for use in the smoothing process. The arrays contain noise estimates, signal estimates, and mean rates for each PDI. The file is MSXPDP42.TXT. 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 provider-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 (OPPDxx variables), the risk-adjusted rates (RPPDxx variables), the smoothed rates (SPDQxx variables), the expected rates (EPPDxx variables), the counts of outcomes that formed the numerators of the observed rates (TPPDxx variables) and the hospital discharge totals that formed the denominators of the observed rates (PPPDxx variables). When stratifications other than hospital are selected, the RPPDxx variables and their confidence intervals are observed/expected ratios to avoid confounding with risk adjustment 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 OPPDxx observed rates and their TPPDxx and PPPDxx numerator and denominator components, the RPPDxx risk-adjusted rates, the SPPDxx smoothed rates and the EPPDxx expected rates. Each record or row in the text file will correspond to a specification group.
3. A PROC MEANS (with N, NMISS, MIN, MAX, MEAN and SUM) of the provider-level summary records that shows statistics for the observed, risk-adjusted and smoothed rates, and statistics for the counts of outcomes that formed the numerators and denominators of the observed rates. These means will only be generated if the user included a value of 16 for the TYPELVLP parameter of the preceding PDSASP2 program(see Table 5 on page 25).
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 requested with the TYPELVLP parameter of the preceding PDSASP2 program (see Table 5 on page 25). If the user does not wish to generate

this printout, then the global “PRINT” parameter in CONTROL\_PDI.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\_PDI.SAS code.
2. In CONTROL\_PDI.SAS, specify the locations (paths) of the input and output SAS files in the “LIBNAME IN1”, “LIBNAME INP2” and “LIBNAME OUTP3” statements.
3. In CONTROL\_PDI.SAS, specify in the “INFILE1” statement the name of the discharge-level file that you created with the PDSAS1 program.
4. In CONTROL\_PDI.SAS, specify in the “FILENAME COVARP” statement the path (including the file name) for the ASCII text file that contains regression coefficients from a regression that was run on the reference SID dataset. This file is provided as part of the PDI module.
5. In CONTROL\_PDI.SAS, specify in the “FILENAME MSXP” statement the path (including the file name) for the ASCII text file that contains the estimates to smooth the indicator rates. This file is provided as part of the PDI module. The name of the file is MSXPDP42.TXT.
6. In CONTROL\_PDI.SAS, specify in the “OUTFILP3” statement the name that you wish the output file to have.
7. In CONTROL\_PDI.SAS, if you wish to create an ASCII (comma delimited) text file that you can import into a spreadsheet, then set the TEXTP3 parameter to 1. Otherwise, the default value of zero for TEXTA3 will skip the code that would generate this file.
8. In CONTROL\_PDI.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 PDTEXTP3” statement.
9. If discharges are weighted, in CONTROL\_PDI.SAS replace “One” in the “DISCWT” statement with the name of the variable in your dataset that contains the weight value for each discharge.

## **8.7 Program 7: Run PDSASA2**

The PDSASA2 program calculates the observed or raw rates for the area-level Pediatric Quality Indicators, using the data derived in a previous step (PDSAS1). These observed rates are stratified by combinations of area, sex, age and race categories. The program first totals the seven Area indicator flags created by the PDSAS1 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 PDSAS1. This dataset is a discharge-level file that contains inpatient records with input variables, stratifiers, and the six new flag indicators for the PDI Area outcomes of interest.
2. A text file (POP95T11.TXT) with Census area residential populations stratified by area, age, sex, and ethnicity categories. The file is currently provided on the website as a separate download along with the PDI module software. The user should set POPYEAR to 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 PDFMT program.

Output:

1. SAS dataset with summary records that contain observed rates (OAPDxx variables where xx refers to the indicator number), the counts of outcomes that formed the numerators of the rates (TAPDxx variables) and the residential population totals that formed the denominators of the observed rates (PAPDxx variables). The output file has records summarized to the various combinations of stratifiers specified in the TYPELVLA parameter that is described in item number 5 in “Changes” 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 OAPDxx observed rates, the TAPDxx counts of outcomes that formed the numerators of the rates, and the PAPDxx residential population totals. The different records/rows in the text file will correspond to the different TYPELVLA 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 OAPDxx observed rates, the TAPDxx counts of outcomes that formed the numerators of the rates, and the PAPDxx residential populations totals. These means will only be generated if the user included a value of 8 for the TYPELVLA parameter discussed in item number 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 TYPELVLA parameter discussed in item number 5 in the “Changes” section below. If the user does not wish to generate this

printout, then the global “PRINT” parameter in CONTROL\_PDI.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\_PDI.SAS code.
2. In CONTROL\_PDI.SAS, specify the locations (paths) of the input and output SAS files in the “LIBNAME IN1” and “LIBNAME OUTA2” statements.
3. In CONTROL\_PDI.SAS, specify the names of the input and output files in the “INFILE1” and “OUTFILEA2” statements. The name that you specify with INFILE1 should be for the file that you created using the PDSAS1 program.
4. In CONTROL\_PDI.SAS, specify the path (including the file name) for the ASCII population text file in the “FILENAME POPFILE” statement. The file POP95T11.TXT is currently provided along with the PDI module software. The user should set POPYEAR to the year that best matches the user’s discharge data file.
5. In CONTROL\_PDI.SAS, specify the levels of aggregation for the observed rates. This is done by specifying numbers between 0 and 15 in the “%LET TYPELVLA =” statement. Each number from 0 to 15 corresponds to a different level or combination of levels as shown in Table 6. The default values of 0 (overall) and 8 (area) will provide an overall total and area-level totals.

 *If TYPELVLA includes the value 0 (overall), 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 PDSAS1 program.*

**Table 6. PDSASA2 aggregation possibilities**

TYPELVLA	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 (overall) and 8 (are) TYPELVLA are default values for the level of computation in the software.

For example, if you are using a state inpatient hospital database, you might specify the following TYPELVLA 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\_PDI.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 the code that would generate this file.
7. In CONTROL\_PDI.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 PDTEXTA2” statement.

## 8.8 Program 8: Run PDSASA3

The PDSASA3 program calculates age and sex risk-adjusted rates for each area-based PDI (overall rates and rates by area), and then calculates smoothed rates.

### Input:

1. SAS format library (FMMLIB) created from the PDFMT program.
2. The discharge-level file that was created with the PDSAS1 program.
3. The SAS dataset with summary records that was created with the PDSASA2 program.
4. A text file with Census area residential populations (stratified by area, sex and discrete age categories). One file is currently provided along with the PDI module software. The file is POP95T11.TXT. The user should set POPYEAR to the year 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. Two text files containing regression coefficients from a regression that was run on the reference SID dataset. These coefficients will be used in the risk adjustment process. COVPDA42.TXT should be used with risk adjustment for age and gender only. CVPDA42A.TXT should be used with adjustment for age, gender, and socioeconomic status (SES). These files are provided to you as part of the PDI module and do not need to be converted to a SAS file for use with the software.
6. Two text files containing two arrays for use in the smoothing process. The arrays contain noise estimates, signal estimates, and mean area rates for each PDI.

MSXPDA42.TXT should be used with adjustment for age and gender only.  
MXPDA42A.TXT should be used with adjustment for age, gender and SES. These files are provided to you as part of the PDI module and do 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 (OAPDxx variables), the risk-adjusted rates (RAPDxx variables), the smoothed rates (SAPDxx variables), the counts of outcomes that formed the numerators of the observed rates (TAPDxx variables) and the residential population totals that formed the denominators of the observed rates (PAPDxx variables). When stratifications other than area are selected, the RAPDxx variables and their confidence intervals are observed/expected ratios to avoid confounding with risk-adjustment 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 OAPDxx observed rates and their TAPDxx and PAPDxx components, the RAPDxx risk-adjusted rates, and the SAPDxx smoothed rates. Each record or row in the text file will correspond to a different area.
3. Output to hardcopy are PROC MEANS (with N, NMISS, MIN, MAX, MEAN, and SUM) of risk-adjusted and smoothed rates in two intermediate work files and of all of the numeric variables in the final output data file. The printout also contains a PROC CONTENTS of the output SAS dataset. A PROC PRINT of the output dataset is provided only if the “PRINT” parameter in CONTROL\_PDI.SAS code is set to one.
4. A PROC MEANS (with N, NMISS, MIN, MAX, MEAN, and SUM) of the area-level summary records that shows statistics for the OAPDxx observed rates and their TAPDxx and PAPDxx components, the RAPDxx risk-adjusted rates, and the SAPDxx smoothed rates.
5. A PROC CONTENTS of the output SAS summary dataset is output to hardcopy.
6. 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\_PDI.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\_PDI.SAS code.
2. In CONTROL\_PDI.SAS, specify the locations (paths) of the input and output SAS files in the “LIBNAME IN1”, “LIBNAME INA2” and “LIBNAME OUTA3” statements.

3. In CONTROL\_PDI.SAS, specify in the “INFILE1” statement the name of the discharge-level file that you created with the PDSAS1 program.
4. In CONTROL\_PDI.SAS, specify the path (including the file name) for the ASCII population text file in the “FILENAME POPFILE” statement. The file name is POP95T11. These files are provided along with the software. The file used should be the one that is closest to the year associated with your particular data. These files contain Census residential estimates by area, gender and discrete age categories.
5. In CONTROL\_PDI.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. To risk adjustment by age and gender, use COVPDA42.TXT. To risk adjustment by SES, use CVPDA42A.TXT.
6. In CONTROL\_PDI.SAS, specify in the “FILENAME MSXA” statement the path (including the file name) for the ASCII text file that contains the estimates to smooth the indicator rates. For age and gender risk adjustment use MSXPDA42.TXT. For age, gender and SES risk adjustment use MXPDA42A.TXT.
7. In CONTROL\_PDI.SAS, specify in the “INFILEA2” statement the name of the summary file that you created with the PDSASA2 program.
8. In CONTROL\_PDI.SAS, specify in the “OUTFILA3” statement the name that you wish the output file to have.
9. In CONTROL\_PDI.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 the code that would generate this file.
10. In CONTROL\_PDI.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 PDTEXTA3” statement.

## 9.0 Reviewing the Printed Output

This section contains tips for reviewing some of the printed output from the PDI module (also see [Appendix D](#) for the output data dictionary). 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.



*The final output of the PDI module is a PROC PRINT from PDSASP2 (for Provider-level Indicators) and PDSASA2 (for Area-level Indicators). All interim printouts are for checking and trouble-shooting.*

### 9.1 PDSAS1

The initial printout from the PDSAS1 program contains **PROC MEANS** output for all of the numeric variables (including the comorbidity indicators) in the output discharge-level dataset. It will contain information for the newly constructed TAPDxx and TPPDxx flag variables that will later form the numerators for the indicator rates. For each TAPDxx and TPPDxx 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 fiveArea Indicators (TAPD14-TAPD18, TAPD90-TAPD92), the MEAN, MINIMUM and MAXIMUM will usually be one since the flag variables have either been set to missing (‘.’) or to a value of one. The Provider Indicators (TPPD01-TPPD01,) will have a value of one, zero, or missing (‘.’). It also includes variables indentifying present on admission (POA) and conditions for provider based outcomes of interest (QPPD01-QPPD01). For these indicators, a value of zero was assigned to the TPPDxx flag if a particular observation was part of the population at risk for a particular indicator but did not have the particular outcome of interest. For example, TPPD01 = 1 indicates a patient who had an accidental puncture or laceration, while TPPD13 = 0 identifies a patient who had a medical or surgical MS-DRG but no accidental puncture or laceration.
- For the Provider Indicators, the MEANs will contain a close approximation of the eventual overall observed indicator rates. The values will change slightly after PDSASP2 has applied additional parts of the indicator definitions.
- N lists the number of observations in the dataset with non-missing values. For the seven Area Indicators, N for TAPDxx will be the same as the SUM. For the Provider Indicators, N will contain the denominator for the observed indicator rate.

The PDSAS1 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 PDSASP2

The purpose of this printout is to provide the observed rates for all Provider-level Indicators. The default printout shows results by provider, but other levels of output (as shown in Table 5)

can be specified as well. Note that risk-adjusted and smoothed rates are shown in the output from the next program.

The printout from the PDSASP2 program contains **PROC MEANS** on the provider-level summary records for all of the numeric variables in the output summary dataset. It will contain information for the newly constructed OPPDxx rates, the PPPDxx denominators and the TPPDxx numerators.

- The `_TYPE_` variable described in the MEANS table identifies the stratification level for the records in the output data set. The `_TYPE_` variable corresponds to the TYPELVLP parameter values that were specified (see item number 5 in “Changes” under the PDSASP2 program on page 32). In this case, `_TYPE_` always assumes the value of 64, since only the provider-level records are selected.
- The N statistic for `_TYPE_` contains the number of records in the output summary data set. A TPPDxx 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 providers. Similarly, a PPPDxx denominator variable with a lower value for N than `_TYPE_` indicates that for some providers, there were no hospital discharges with the outcome of interest.
- The MINIMUM value for the TPPD01-TPPD13 numerators will 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 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 hospital means; i.e., all hospitals are given equal weight, regardless of the number of discharges in the different hospitals.*

- The SUMs for the counter variables TPPDxx and PPPDxx 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\_PDI.SAS program is set to 1, the subsequent **PROC PRINT** output contains a complete dump of the output summary file. Listed for each record are the stratification values, the PPPDxx population denominators for the 13 indicators, the TPPDxx outcome numerators for the 13 indicators and the 13 OPPDxx observed rates.

### 9.3 PDSASP3

 *The **PROC PRINT** at the end of this program provides your **final output**. (This printout appears if the “PRINT” parameter in the CONTROL\_PDI.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 PDI number.

**Table 7. Provider level printed output**

Column Heading	Description
Obs	Observation 1 (with HOSPID = ".") is the overall average for the entire dataset (TYPELVLP = 0). The remaining observations are individual providers, or hospitals (TYPELVLP = 16).
HOSPID	Identifier for each provider in the dataset.
AGECAT	Age stratification category - See Table 6 for TYPELVLP settings.
SEXCAT	Gender stratification category - See Table 6 for TYPELVLP settings.
PAYCAT	Payer stratification category - See Table 6 for TYPELVLP settings.
RACECAT	Race stratification category - See Table 6 for TYPELVLP settings.
TPPDxx	Number of cases in the numerator.
PPPDxx	Number of cases in the denominator (population at risk).
OPPDxx	Observed (raw) rate - Numerator divided by Denominator.
EPPDxx	Expected rate - Rates the provider would have if it performed the same as the reference population given the provider's actual case-mix (e.g., age, gender, modified DRG and comorbidities)
RPPDxx	Risk-adjusted rate - Accounts for the difference between the case-mix of the reference population and the provider's case-mix. When stratifications other than hospital are selected, the RPPDxx variables and their confidence intervals are observed/expected ratios to avoid confounding with risk adjustment variables
LPPDxx	Lower limit of confidence interval of risk-adjusted rate.
UPPDxx	Upper limit of confidence interval of risk-adjusted rate.
SPPDxx	Smoothed rate - Rates with removal of fluxuation over time.
XPPDxx	Standard error of smoothed rate.

Observation 1 (with HOSPID = .) is the overall average for the entire dataset (TYPELVLP = 0). The remaining observations are individual providers, or hospitals (TYPELVLP = 16). Data will appear under the column headers, AGECAT, SEXCAT, PAYCAT, and RACECAT, if you specify these levels of aggregation (shown in Table 5).

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

Multiply the rates by 100 to express them as a percentage (e.g.,  $0.0338 * 100 = 3.4\%$  mortality for pediatric heart surgery).

The MEANS table is generated from the permanent provider-based output file and is provided for trouble-shooting and understanding your data, but not for reporting. The means are provided just for the provider level summary records.

The N for all but the TPPDxx counter variables should contain the number of providers in the user's database.

The MEANS, MINIMUMs, and MAXIMUMs have their normal meaning and provide comparisons among the different hospitals or providers in your 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 in the PROC PRINT labeled "FINAL OUTPUT." The MEANS given here are means of

*hospital means; i.e., all hospitals are given equal weight, regardless of the number of discharges in the different hospitals.*



If the “PRINT” parameter in the CONTROL\_PDI.SAS program is set to one, at the end of the printout is a **PROC PRINT** on the final provider-based output file. Listed for each level of stratification are the stratifiers, the numerator and denominator totals for the observed rates, and the observed, risk-adjusted, and smoothed rates.

## 9.4 PDSASA2

The purpose of this printout is to provide the observed rates for all Area-level Indicators. The default printout shows overall results and results by area, but other levels of output (as shown in Table 7) can be specified as well.

- The printout from the PDSASA2 program contains **PROC MEANS** for all of the numeric variables in the output summary dataset. It will contain information for the newly constructed OAPDxx rates, the PAPDxx denominators, and the TAPDxx 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 TYPELVLA parameter values that were specified (see item 5 in “[Changes](#)” under the PDSASA2 program on page 32). 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 TAPDxx numerator variable with a lower value for N than `_TYPE_` indicates that there were no outcomes of interest for some of the areas
- The MINIMUM value for the TAPDxx numerators 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 the area means; i.e., all areas are given equal weight, regardless of the population in the different areas.*

- The SUMs for the counter variables TAPDxx and PAPDxx 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\_PDI.SAS program is set to 1, the subsequent **PROC PRINT** output contains a complete dump of the output summary file. Listed for each record are the stratification values, the PAPDxx population denominators for the seven indicators, the TAPDxx outcome numerators for the seven indicators, and the seven OAPDxx observed rates.

## 9.5 PDSASA3

 *The PROC PRINT at the end of this program provides your final output. (This printout appears if the “PRINT” parameter in the CONTROL\_PDI.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 PDI number.

**Table 8. Area level printed output**

Column Heading	Description
Obs	Observation 1 is the overall average for the entire dataset (TYPELVLA = 0). The remaining observations are individual areas (TYPELVLA = 8).
Area	Identifier in the dataset for area.
TAPDxx	Number of cases in the numerator.
PAPDxx	Number of cases in the denominator (population at risk).
OAPDxx	Observed (raw) rate - Numerator divided by Denominator.
RAPDxx	Risk-adjusted rate - Accounts for the difference between the case-mix of the reference population and the area's case-mix. When stratifications other than area are selected, the RAPDxx variables and their confidence intervals are observed/expected ratios to avoid confounding with risk adjustment variables
SAPDxx	Smoothed rate - Rates with removal of fluxuation over time.

Observation 1 (with AREA = *blank*) is the overall average for the entire dataset (TYPELVL = 0). The remaining observations are individual areas (TYPELVL = 8).

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

Multiply the utilization rates by 100,000 to express them as a rate per 100,000 population (e.g.,  $0.000494 * 100,000 = 49.4$  pediatric heart surgeries 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 TAPDxx counter variables should contain the number of different areas (Metro Areas and counties) in your database.

The MEANSs, 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 TAPDxx and PAPDxx variables) are associated with specific areas, and therefore these MAXIMUMs may not match those in the prior PDSAS2 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 TAPDxx and PAPDxx variables) yield the overall database totals.

## 10.0 User Support and Information

Technical assistance is available, through an electronic user support system monitored by the QI support team, to support users in their application of the PDI software. The same e-mail address may be used to communicate to AHRQ any suggestions for PDI enhancements, general questions, and any QI related comments you have. AHRQ welcomes your feedback.

The Internet address for user support and feedback is: [support@qualityindicators.ahrq.gov](mailto:support@qualityindicators.ahrq.gov)

AHRQ also offer a listserv to keep you informed on the Quality Indicators (QIs). The listserv is used to announce any QI changes or updates, new tools and resources, and to distribute other QI related information. This is a free service. Simply follow the process described below to begin receiving important QI information. 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](mailto: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. You should receive an automatic response e-mail message confirming receipt of your e-mail. 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 Pediatric Quality Indicators.

### Pediatric Quality Indicators Version 4.3 Documents and Software

Available at [http://www.qualityindicators.ahrq.gov/modules/pdi\\_resources.aspx](http://www.qualityindicators.ahrq.gov/modules/pdi_resources.aspx) and <http://www.qualityindicators.ahrq.gov/software/default.aspx>

Title	Description
<i>PDI User Guide: Individual Measures</i>	The user guide provides a technical overview of the PDI individual measures, including literature and empirical evidence.
<i>PDI User Guide: Composite Measures</i>	The user guide provides a technical overview of the PDI composite measures.
<i>PDI Technical Specifications</i>	Provides detailed definitions of each PDI, 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.
<i>PDI Comparative Data</i>	This document provides the provider rate, provider standard deviation, area rate, and population rate, as appropriate, for each indicator.
<i>Fiscal year 2011 Coding Changes</i>	This document summarizes the changes to the indicator definitions resulting from FY 2011 changes to ICD-9-CM coding and DRG changes. These changes will only affect data from FY 2011 (October 1, 2010) or later.
<i>Log of Changes to PDI Documents and Software</i>	The Change Log document provides a cumulative summary of all changes to the PDI software, software documentation, and other documents made since the release of version 3.0 of the software in February 2006. Changes to indicator specifications that were not a result of new ICD-9-CM and DRG codes, are also described in the Change Log.
<i>PDI Risk Adjustment Coefficient Tables</i>	Tables for each PDI provide the stratification and coefficients used to calculate the risk-adjusted rate for each strata.
<i>SAS® PDI Software Documentation</i>	This software documentation provides detailed instructions on how to use the SAS ® version of the PDI software including data preparation, calculation of the PDI rates, and interpretation of output.
<i>Measures of Pediatric Health Care Quality Based on Hospital Administrative Data, The Pediatric Quality Indicators.</i>	Describes how the PDI were developed and provides detailed evidence for each indicator. Four appendices are also available, describing literature methods, clinical review panels, peer reviewers, and a panel packet sample.

Title	Description
SAS® PDI Software	Requires the SAS® statistical program distributed by the SAS Institute, Inc. The company may be contacted directly regarding the licensing of its products: <a href="http://www.sas.com">http://www.sas.com</a>
<i>Version 4.3 Listserv announcement</i>	This document announces the release of the Version 4.3 software and documentation, provides a summary of changes and links to relevant pages.

### **AHRQ QI Windows Application**

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

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

### **Additional Documents**

The following documents are available within the “Archive” section of the AHRQ QI Web page:

<http://www.qualityindicators.ahrq.gov/Archive/default.aspx#pdi>

- *Refinement of the HCUP Quality Indicators (Technical Review), May 2001*
- *Refinement of the HCUP Quality Indicators (Summary), May 2001*

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

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

- *AHRQ QI Measure Development, Implementation, Maintenance and Retirement*
- *ICD-9-CM to ICD-10-CM/PCS Conversion of AHRQ QI, March 2011*
- *AHRQ QI Empirical Methods, May 2011*
- *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 Existing National Frameworks*
- *List of AHRQ QI Endorsed by the National Quality Forum*

The following documents can be viewed or downloaded from the AHRQ QI News Web page:

<http://www.qualityindicators.ahrq.gov/news/default.aspx> and  
<http://www.qualityindicators.ahrq.gov/Archive/News.aspx>

- *February 2007 Newsletter - Year highlights and future plans*
- *February 2006 Newsletter – Year highlights and future plans*
- *2006 Area Level Indicator Changes*
- *Considerations in Public Reporting for the AHRQ QI*
- *June 2005 Newsletter - Contains the article, “Using Different Types of QI Rates”*

### **Other Tools and Information**

Before running the PDI software, users who intend to calculate risk-adjusted rates using PDSASP3 will need to download and run the program that creates the comorbidity variables in the user’s data file. The program is available at:

<http://www.hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp>

Area indicators 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>

## Appendix B: Using Different Types of QI Rates

When should you use the observed, expected, risk adjusted, and/or smoothed rates generated by the AHRQ QI software? Here are some guidelines.

If the user's primary interest is to identify cases for further follow-up and quality improvement, then the observed rate would help to identify them. The observed rate is the raw rate generated by the QI software from the data the user provided. Areas for improvement can be identified by the magnitude of the observed rate compared to available benchmarks and/or by the number of patients impacted.

Additional breakdowns by the default patient characteristics used in stratified rates (e.g., age, gender or payer) can further identify the target population. Target populations can also be identified by user-defined patient characteristics supplemented to the case/discharge level flags. Trend data can be used to measure change in the rate over time.

Another approach to identify areas to focus on is to compare the observed and expected rates. The expected rate is the rate the provider would have if it performed the same as the reference population given the provider's actual case-mix (e.g., age, gender, DRG and comorbidity categories).

If the observed rate is higher than the expected rate (i.e., the ratio of observed/expected is greater than 1.0, or observed minus expected is positive), then the implication is that the provider performed worse than the reference population for that particular indicator. Users may want to focus on these indicators for quality improvement.

If the observed rate is lower than the expected rate (i.e., the ratio of observed/expected is less than 1.0, or observed minus expected is negative), then the implication is that the provider performed better than the reference population. Users may want to focus on these indicators for identifying best practices.

Users can also compare the expected rate to the population rate reported in the Comparative Data document to determine how their case-mix compares to the reference population. The population rate refers to the overall rate for the reference population. The reference population is defined in the Comparative Data Document. If the population rate is higher than the expected rate, then the provider's case-mix is less severe than the reference population. If the population rate is lower than the expected rate, then the provider's case-mix is more severe than the reference population.

We use this difference between the population rate and the expected rate to "adjust" the observed rate to account for the difference between the case-mix of the reference population and the provider's case-mix, which is the provider's risk-adjusted rate.

If the provider has a less severe case-mix, then the adjustment is positive (population rate > expected rate) and the risk-adjusted rate is higher than the observed rate. If the provider has a more severe case-mix, then the adjustment is negative (population rate < expected rate) and the risk-adjusted rate is lower than the observed rate. The risk-adjusted rate is the rate the provider would have if it had the same case-mix as the reference population given the provider's actual performance.

Finally, users can compare the risk-adjusted rate to the smoothed or "reliability-adjusted" rate to determine whether this difference between the risk-adjusted rate and reference population rate is likely to remain in the next measurement period. Smoothed rates are weighted averages of the population rate and the risk-adjusted rate, where the weight reflects the reliability of the provider's risk-adjusted rate.

A ratio of (smoothed rate - population rate) / (risk-adjusted rate - population rate) greater than 0.80 suggests that the difference is likely to persist (whether the difference is positive or negative). A ratio less than 0.80 suggests that the difference may be due in part to random differences in patient characteristics (patient characteristics that are not observed and controlled for in the risk-adjustment model). In general, users may want to focus on areas where the differences are more likely to persist.

## Appendix C: PDI Input Data Dictionary

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 QI programs, but should be present to facilitate possible exploration; allows user to link the records from the PDSAS1.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 the analysis.
AGEDAY	Age in Days (coded only when the age in years is less than 1)	Numeric	Age in days	If this data element is missing alternative logic will sometimes apply.
RACE	Race of Patient.	Numeric 1 2 3 4 5 6	White Black Hispanic Asian or Pacific Island Native American Other	The values 1 through 5 are used directly in the QI software. All other ethnicity codes are mapped to an 'other' category.
SEX	Sex of Patient	Numeric 1 2	Male Female	If this data element is missing the discharge record will be excluded from the analysis.
PAY1	Expected Primary Payer	Numeric 1 2 3 4 5 6	Medicare Medicaid Private, incl. HMO Self-pay No charge Other	The values 1, 2, 3, and 4 are used directly in the QI software. All other payer codes are mapped to an 'other' category. This data element is used to stratify only the Provider-level PDI (#01 to #13).
PAY2 (SAS Only)	Expected secondary payer	Numeric 1 2 3 4 5 6	Medicare Medicaid Private, incl. HMO Self-pay No charge Other	Use is optional but must be present in data (set to missing if data element not available).
PSTCO	Location of Patient Residence or Hospital Location (FIPS State/County Code)	Numeric ssccc	Modified Federal Information Processing Standards State/county code <sup>3</sup>	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. May also be selected as a stratifier (see Table 6. PDSASP2 Aggregation Possibilities).

<sup>3</sup> 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 U.S. Census Bureau 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. A list is available at <http://www.census.gov/popest/geographic/codes02.pdf>

Variable name	Description	Format	Value description	Comments
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 in the hospital	The values 2 and 20 are referenced by the QI code (to identify transfers to another short-term hospital and patients who died in the hospital).
MORT30 (SAS Only)	30-Day Mortality Indicator	Numeric	0 – not died within 30-days from admission 1 – died within 30- days from admission	Use is optional but must be present in data (set to missing if data element not available).
DNR (SAS Only)	Do not resuscitate	Numeric 0 1	No "Do Not Resuscitate" order "Do Not Resuscitate" order	Use is optional but must be present in data (set to missing if data element not available).
ATYPE	Admission Type	Numeric 1 2 3 4 5 6	Emergency Urgent Elective Newborn Trauma Center <sup>4</sup> Other	The values 3 and 4 are referenced by the PDI code (to identify elective surgeries). PDI 8 and 9 will be affected if ATYPE values are missing.
ASOURCE	Admission Source	Numeric 1 2 3 4 5	ER Another hospital Another facility including LTC Court/law enforcement Routine/birth/other	The values 2 and 3 are referenced by the PDI code (to identify transfers from another hospital or facility).
LOS	Length of Stay	Numeric	Number of days from admission to discharge	
DRG	Diagnosis Related Group	Numeric	DRG from federal (CMS) Grouper	
DRGVER	Diagnosis Related Group Version	Numeric	DRG Version of federal (CMS) DRG Grouper	e.g. Version 28 for FY2011
MDC	Major Diagnostic Category	Numeric	MDC from federal (CMS) grouper	
DISCWT	Weight to assign to discharge	Numeric	Format as "nn.nnn"	Cases with missing values are assigned a value of 1. Assigning a weight of zero has the effect of removing the case from the analysis.

<sup>4</sup> 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
DX1 – DX30	ICD-9-CM Diagnosis Codes. DX1 is the principal diagnosis. DX2- DX30 are secondary diagnoses.	String, 3, 4 or 5 characters (do not include decimal point)	Diagnosis codes	Users must specify the number of diagnoses in the parameter &NDX in the CONTROL_PDI.SAS program to reflect the number of diagnoses.
DXPOA1- DXPOA30-	Present on admission indicator for each diagnosis code	String	1 = present at the time of inpatient admission	These data will be used in risk adjustment and smoothing if the alternative parameter files of risk-adjustment covariates and population rates is specified.
			('Y','W','E','1')	
			0 = not present at the time of inpatient admission	
			('N','U','0','')	
NDX	Count of the Number of Diagnoses (DX) on the Discharge Record.	Numeric	Count of diagnoses	Not used by the SAS program. If not present in the input data file it is not necessary to create a dummy variable. Definition is: 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.
			(principal and all secondary diagnoses)	
PR1 – PR30	ICD-9-CM Procedure Codes. PR1 is the principal procedure, PR2- PR30 are secondary procedures.	String, 3 or 4 characters (do not include decimal point)"	Procedure codes	Users with more or fewer secondary procedures should modify the parameter &NPR in CONTROL_PDI.SAS to reflect the number of procedures.
NPR	Count of the Number of Procedures (PR) on the Discharge Record.	Numeric	Count of procedures	Not used by the SAS program. If not present in the input data file it is not necessary to create a dummy variable. Definition is: 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.
			(principal and all secondary procedures)	
POINTOFORIGINUB04	Point of Origin	Numeric	4 Transfer from a hospital 5 Transfer from a skilled Nursing Facility (SNF) or Intermediate Care Facility (ICF)	Only these values are used by the QI programs.
			6 Transfer from another health care facility	
			IF ATYPE = 4	
			5 Born inside this hospital	
			6 Born outside of this hospital	

Variable name	Description	Format	Value description	Comments
PRDAY1 – PRDAY30	Days from Admission to Procedure. PR1 is the principal procedure, PR2-PR30 are secondary procedures.	Numeric	Days from admission to procedure <sup>5</sup>	Necessary variables if the user sets the &PRDAY parameter in CONTROL_PDI.SAS to equal one. In this case, it is expected that the number of PRDAY variables agrees with the number of Procedure codes present. (See note below for method of calculation). These variables are not needed if the user sets the &PRDAY parameter in CONTROL_PDI.SAS to equal zero, indicating that there is no procedure day information available.
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 2011.	
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	

\* See <http://hcup-us.ahrq.gov/toolssoftware/comorbidity/comorbidity.jsp>.

Note: If e-codes are separated from secondary diagnoses in the input data file, the variable should be renamed and included as a secondary diagnosis variable (e.g., e-codes would be labeled as DX10 in a data file where the last secondary DX field is DX9).

<sup>5</sup> Variables PRDAY1 to PRDAY30 are defined as the number of days from date of admission to date of procedure for each procedure. For example, if the patient was admitted on June 15<sup>th</sup> and had two procedures - the principal procedure on June 15<sup>th</sup> and a second procedure on June 18<sup>th</sup> – then the value of PRDAY1 would be zero (0) and the value of PRDAY2 would be three (3). For more information, consult the HCUP data documentation at <http://www.hcup-us.ahrq.gov/db/vars/prdayn/nisnote.jsp>.

## Appendix D: SAS PDI Output Data Dictionary

### SAS PDI Provider Output

Column Heading	Description	Format	Comments
Obs	Observation 1 (with HOSPID = ".") is the overall average for the entire dataset (TYPELVLP = 0). The remaining observations are individual providers, or hospitals (TYPELVLP = 16).	Numeric	
HOSPID	Identifier for each provider in the dataset.	Text	Up to 12 characters
AGECAT	Age stratification category	Numeric range	
SEXCAT	Gender stratification category	Numeric	1 = male; 2 = female
PAYCAT	Payer stratification category	Numeric	1 = Medicare ; 2 = Medicaid ; 3 = Private, incl. HMO ; 4 = Self-pay ; 5 = No charge ; 6 = Other
RACECAT	Race stratification category	Numeric	1 = White ; 2 = Black ; 3 = Hispanic ; 4 = Asian or Pacific Island ; 5 = Native American ; 6 = Other
TPPDxx	The number of discharge records included in the numerator (outcome of interest) as defined for the indicator	Numeric	
PPPDxx	The number of discharge records included in the denominator (population at risk) as defined for the indicator	Numeric	
OPPDxx	The rate (observed numerator / observed denominator) as defined for the indicator	Numeric	
EPPDxx	Rate calculated by assuming an "average" performance for each patient group based on the reference population, but with the provider's actual case-mix	Numeric	
RPPDxx	The estimated rate calculated by adjusting to an "average" case-mix	Numeric	When stratifications other than hospital or area are selected, the RPPDxx variables and their confidence intervals are observed/expected ratios to avoid confounding with risk adjustment variables
LPPDxx	The lower confidence bound of the risk adjusted rate	Numeric	

Column Heading	Description	Format	Comments
UPPDxx	The upper confidence bound of the risk adjusted rate	Numeric	
SPPDxx	The smoothed rate calculated using multivariate signal extraction (MSX)	Numeric	
XPPDxx	Standard error of smoothed rate	Numeric	
VPPDxx	Variance for respective indicators	Numeric	

### SAS PDI Area Output

Column Heading	Description	Format	Comments
Obs	Observation 1 (with HOSPID = ".") is the overall average for the entire dataset (TYPELVLP = 0). The remaining observations are individual providers, or hospitals (TYPELVLP = 16).	Numeric	
AREA	Identifier in the dataset for area	Text	
AGECAT	Age stratification category	Numeric range	
SEXCAT	Gender stratification category	Numeric	
PAYCAT	Payer stratification category	Numeric	
RACECAT	Race stratification category	Numeric	
TAPDxx	The number of discharge records included in the numerator (outcome of interest) as defined for the indicator	Numeric	
PAPDxx	The number of discharge records included in the denominator (population at risk) as defined for the indicator	Numeric	
OAPDxx	The rate (observed numerator / observed denominator) as defined for the indicator	Numeric	
RAPDxx	The estimated rate calculated by adjusting to an "average" case-mix	Numeric	When stratifications other than hospital or area are selected, the RAPDxx variables and their confidence intervals are observed/expected ratios to avoid confounding with risk adjustment variables
SAPDxx	The smoothed rate calculated using multivariate signal extraction (MSX)	Numeric	