

HEALTHCARE COST AND UTILIZATION PROJECT — HCUP
A FEDERAL-STATE-INDUSTRY PARTNERSHIP IN HEALTH DATA
Sponsored by the Agency for Healthcare Research and Quality

INTRODUCTION TO
THE HCUP KIDS' INPATIENT DATABASE (KID)
2012

These pages provide only an introduction to the KID package.

**For full documentation and notification of changes,
visit the HCUP User Support (HCUP-US) Web site at
<http://www.hcup-us.ahrq.gov>.**

Issued July 2014
Updated Nov. 2015

Agency for Healthcare Research and Quality
Healthcare Cost and Utilization Project (HCUP)

Phone: (866) 290-HCUP (4287)
Email: hcup@ahrq.gov
Web site: <http://www.hcup-us.ahrq.gov>

KID Data and Documentation Distributed by:

HCUP Central Distributor
Phone: (866) 556-4287 (toll-free)
Fax: (866) 792-5313
Email: HCUPDistributor@ahrq.gov

Table of Contents

SUMMARY OF DATA USE RESTRICTIONS.....	1
HCUP CONTACT INFORMATION.....	2
WHAT'S NEW IN THE 2012 KIDS' INPATIENT DATABASE (KID)?	3
UNDERSTANDING THE KID.....	3
ABSTRACT.....	4
INTRODUCTION TO THE HCUP KIDS' INPATIENT DATABASE (KID)	6
Overview of KID Data.....	6
<i>KID Data Sources, Hospitals, and Inpatient Stays</i>	7
<i>State-Specific Restrictions</i>	7
<i>Contents of KID</i>	8
<i>KID Data Elements</i>	9
Getting Started.....	9
<i>Copying and Decompressing the ASCII Files</i>	9
<i>Downloading and Running the Load Programs</i>	9
<i>KID Documentation</i>	10
<i>HCUP On-Line Tutorials</i>	10
HOW TO USE THE KID FOR DATA ANALYSIS.....	11
Calculating National Estimates.....	11
Studying Trends.....	12
Choosing Data Elements for Analysis	12
Hospital-Level Data Elements	13
ICD-9-CM Diagnosis and Procedure Codes	13
Missing Values.....	13
Variance Calculations.....	14
Computer Software for Variance Calculations	15
SAMPLING OF DISCHARGES	16
Sampling of Discharges Included in the KID.....	16
<i>The KID Hospital Universe</i>	16
<i>Hospital Merges, Splits, and Closures</i>	17
<i>Stratification Data Elements</i>	17
Hospital Sampling Frame	19
Hospital Sample Design	19
<i>Design Considerations</i>	19
<i>Changes to Sampling and Weighting Strategy Beginning with the 2000 KID</i>	20
<i>Sampling Procedure</i>	21
SAMPLE WEIGHTS.....	21
Discharge Weights	21

Weight Data Elements.....	23
THE FINAL KID SAMPLE	23
APPENDIX I: TABLES AND FIGURES	I-1
Table 1. Data Sources for the 2012 KID	I-1
Table 2. Summary of KID States, Hospitals, and Inpatient Stays, 1997-2012.....	I-3
Figure 1. KID States, by Region, 2012	I-4
Table 3. Summary of KID Releases	I-5
Table 4. KID Related Reports and Database Documentation Available on HCUP-US.....	I-6
Table 5. Bed Size Categories, by Region	I-7
Table 6: Number of AHA, HCUP SID, and KID Hospitals, by Region (the KID Stratifier) and by Census Division, 2012	I-8
Table 7. Number of Hospitals and Discharges in the AHA Universe, SID, and KID, by Hospital Type, 2012.....	I-9
Table 8. 2012 KID Discharges, by Hospital Type	I-10
Figure 2. Number of Hospitals in the 2012 AHA Universe, SID, and KID, by Region	I-11
Figure 3. Percentage of U.S. Population in 2012 KID States, by Region	I-12
APPENDIX II: STATE-SPECIFIC RESTRICTIONS.....	II-1
APPENDIX III: DATA ELEMENTS	III-1
Table 1. Data Elements in the 2012 KID Inpatient Core File.....	III-1
Table 2. Data Elements in the 2012 KID Hospital File	III-4
Table 3. Data Elements in the 2012 KID Disease Severity Measures Files	III-5
Table 4. Data Elements in the 2012 KID Diagnosis and Procedure Groups Files	III-7
APPENDIX IV: TEACHING HOSPITAL INDICATOR ASSIGNMENT	IV-1

**HCUP KIDS' INPATIENT DATABASE (KID)
SUMMARY OF DATA USE RESTRICTIONS**

******* REMINDER *******

All users of the KID must take the on-line HCUP Data Use Agreement (DUA) training course, and read and sign a Data Use Agreement.†

Authorized users of HCUP data agree to the following restrictions: ‡

- Will not use the data for any purpose other than research or aggregate statistical reporting.
- Will not re-release any data to unauthorized users.
- Will not redistribute HCUP data by posting on any Website or other publically-accessible online repository.
- Will not identify or attempt to identify any individual, including by the use of vulnerability analysis or penetration testing. Methods that could be used to identify individuals directly or indirectly shall not be disclosed or published.
- Will not publish information that could identify individual establishments (e.g., hospitals) and will not contact establishments.
- Will not use the data concerning individual establishments for commercial or competitive purposes involving those establishments, and will not use the data to determine rights, benefits, or privileges of individual establishments.
- Will not use data elements from the proprietary severity adjustment software packages (3M APR-DRGs, HSS APS-DRGs, and Thomson Reuters Disease Staging) for any commercial purpose or to disassemble, decompile, or otherwise reverse engineer the proprietary software.
- Will acknowledge in reports that data from the "Healthcare Cost and Utilization Project (HCUP)" were used, including names of the specific databases used for analysis.
- Will acknowledge that risk of individual identification of persons is increased when observations (i.e., individual discharge records) in any given cell of tabulated data is less than or equal to 10.

Any violation of the limitations in the Data Use Agreement is punishable under Federal law by a fine of up to \$10,000 and up to 5 years in prison. Violations may also be subject to penalties under State statutes.

† The on-line Data Use Agreement training session and the Data Use Agreement are available on the HCUP User Support ([HCUP-US](http://www.hcup-us.ahrq.gov)) Web site at <http://www.hcup-us.ahrq.gov>.
‡ Specific provisions are detailed in the Data Use Agreement for Nationwide Databases.

HCUP CONTACT INFORMATION

All HCUP data users, including data purchasers and collaborators, must complete the online HCUP Data Use Agreement (DUA) Training Tool, and read and sign the HCUP Data Use Agreement. Proof of training completion and signed Data Use Agreements must be submitted to the HCUP Central Distributor as described below.

The on-line DUA training course is available at: http://www.hcup-us.ahrq.gov/tech_assist/dua.jsp.

The HCUP Nationwide Data Use Agreement is available on the AHRQ-sponsored HCUP User Support (HCUP-US) Web site at: <http://www.hcup-us.ahrq.gov>

HCUP Central Distributor

Data purchasers will be required to provide their DUA training completion code and will execute their DUAs electronically as a part of the online ordering process. The DUAs and training certificates for collaborators and others with access to HCUP data should be submitted directly to the HCUP Central Distributor using the contact information below.

The HCUP Central Distributor can also help with questions concerning HCUP database purchases, your current order, training certificate codes, or invoices, if your questions are not covered in the Purchasing FAQs on the HCUP Central Distributor Web site.

Purchasing FAQs:

<https://www.distributor.hcup-us.ahrq.gov/Purchasing-Frequently-Asked-Questions.aspx>

Phone: 866-556-HCUP (4287) (toll free)

Email: HCUPDistributor@AHRQ.gov

Fax: 866-792-5313 (toll free in the United States)

Mailing address:

HCUP Central Distributor

Social & Scientific Systems, Inc.

8757 Georgia Ave, 12th Floor

Silver Spring, MD 20910

HCUP User Support:

Information about the content of the HCUP databases is available on the HCUP User Support (HCUP-US) Web site (<http://www.hcup-us.ahrq.gov>). If you have questions about using the HCUP databases, software tools, supplemental files, and other HCUP products, please review the HCUP Frequently Asked Questions or contact HCUP User Support:

HCUP FAQs:

http://www.hcup-us.ahrq.gov/tech_assist/faq.jsp

Phone: 866-290-HCUP (4287) (toll free)

Email: hcp@ahrq.gov

WHAT'S NEW IN THE 2012 KIDS' INPATIENT DATABASE (KID)?

- The 2012 Kids' Inpatient Database (KID) incorporates the following changes to enhance confidentiality:
 - The 2012 KID eliminates State identifiers, unencrypted hospital identifiers, and other data elements that are not uniformly available across States.
 - HOSP_KID is now the encrypted hospital number. Prior to the 2012 KID, the HCUP hospital identifier (HOSPID) provided the linkage between the KID Inpatient Core files and the Hospital Weights file. Beginning with the 2012 KID, the KID hospital number (HOSP_KID) provides the linkage between the KID Inpatient Core files and the Hospital Weights file. The HOSP_KID values are reassigned each year, so they cannot be used to link hospitals across years.
 - The 2012 KID no longer includes hospital identifiers that can be linked to other hospital-level data (such as the AHA Annual Survey).
 - KID_STRATUM=9999 (or 9998 in 2003) is now the way to identify children's hospitals. Prior to 2012, children's hospitals could be selected using the NACHTYPE data element in the KID. Beginning with 2012 data, NACHTYPE is no longer available, but discharges from freestanding children's hospitals are stratified as children's hospitals (KID_STRATUM=9999 or 9998).
 - HOSP_DIVISION was added to indicate the hospital census division—New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, or Pacific. REGION remains the data element used in post-stratification and represents an aggregation of divisions.

UNDERSTANDING THE KID

This document, Introduction to the KID, 2012, summarizes the content of the KID and describes the development of the KID sample and weights. Cumulative information for all previous years is included to provide a longitudinal view of the database. Important considerations for data analysis are provided along with references to detailed reports. In-depth documentation for the KID is available on the HCUP User Support (HCUP-US) Web site (www.hcup-us.ahrq.gov).

HEALTHCARE COST AND UTILIZATION PROJECT — HCUP
A FEDERAL-STATE-INDUSTRY PARTNERSHIP IN HEALTH DATA
Sponsored by the Agency for Healthcare Research and Quality

**The Agency for Healthcare Research and Quality and
the staff of the Healthcare Cost and Utilization Project (HCUP) thank you for purchasing
the HCUP Kids' Inpatient Database (KID).**

HCUP Kids' Inpatient Database (KID)

ABSTRACT

The Kids' Inpatient Database (KID) is part of the Healthcare Cost and Utilization Project (HCUP), sponsored by the Agency for Healthcare Research and Quality (AHRQ).

The KID is the largest publicly-available all-payer pediatric inpatient care database in the United States, yielding national estimates of hospital inpatient stays by children. The KID is a sample of pediatric discharges from all community, non-rehabilitation hospitals in States participating in HCUP. The target universe includes pediatric discharges from community, non-rehabilitation hospitals in the United States. Pediatric discharges are defined as all discharges where the patient was age 20 or less at admission. See [Table 1](#) in [Appendix I](#) for a list of the statewide data organizations participating in the KID. The number of sample hospitals and discharges by State and year are available in [Table 2](#) in [Appendix I](#).

Inpatient stay records in the KID include clinical and resource use information typically available from discharge abstracts created by hospitals for billing. The KID contains charge information on all patients, regardless of payer, including persons covered by private insurance, Medicaid, Medicare, and the uninsured. The KID's large sample size enables analyses of rare conditions, such as congenital anomalies and uncommon treatments, such as cardiac surgery. It can be used to study a wide range of topics including the economic burden of pediatric conditions, access to services, quality of care and patient safety, and the impact of health policy changes. Discharge weights are provided for calculating national estimates.

- The 2012 Kids' Inpatient Database (KID) incorporates the following changes to enhance confidentiality:
 - The 2012 KID eliminates State identifiers, unencrypted hospital identifiers, and other data elements that are not uniformly available across States.
 - HOSP_KID is now the encrypted hospital number. Prior to the 2012 KID, the HCUP hospital identifier (HOSPID) provided the linkage between the KID Inpatient Core files and the Hospital Weights file. Beginning with the 2012 KID, the KID hospital number (HOSP_KID) provides the linkage between the KID Inpatient Core files and the Hospital Weights file. The HOSP_KID values are reassigned each year, so they cannot be used to link hospitals across years.
 - The 2012 KID no longer includes hospital identifiers that can be linked to other hospital-level data (such as the AHA Annual Survey).

- KID_STRATUM=9999 (or 9998 in 2003) is now the way to identify children's hospitals. Prior to 2012, children's hospitals could be selected using the NACHTYPE data element in the KID. Beginning with 2012 data, NACHTYPE is no longer available, but discharges from freestanding children's hospitals are stratified as children's hospitals (KID_STRATUM=9999 or 9998).
- HOSP_DIVISION was added to indicate the hospital census division—New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain, or Pacific. REGION remains the data element used in post-stratification and represents an aggregation of divisions.

The KID is available every three years beginning with 1997. Periodically, new data elements are added to the KID and some are dropped; see [Appendix III](#) for a summary of data elements and when they are effective.

Access to the KID is open to users who sign Data Use Agreements. Uses are limited to research and aggregate statistical reporting.

For more information on the KID, visit the AHRQ-sponsored HCUP User Support ([HCUP-US](#)) Web site at <http://www.hcup-us.ahrq.gov>.

INTRODUCTION TO THE HCUP KIDS' INPATIENT DATABASE (KID)

Overview of KID Data

The Kids' Inpatient Database (KID) is part of the Healthcare Cost and Utilization Project (HCUP), sponsored by the Agency for Healthcare Research and Quality (AHRQ).

The KID is the largest publicly-available all-payer pediatric inpatient care database in the United States, yielding national estimates of hospital inpatient stays by children. The KID is a sample of pediatric discharges from all community, non-rehabilitation hospitals in States participating in HCUP. See [Table 1](#) in [Appendix I](#) for a list of the statewide data organizations participating in the KID.

Inpatient stay records in the KID include clinical and resource use information typically available from discharge abstracts created by hospitals for billing. The KID contains charge information on all patients, regardless of payer, including persons covered by private insurance, Medicaid, Medicare, and the uninsured. The KID's large sample size enables analyses of rare conditions, such as congenital anomalies and uncommon treatments, such as cardiac surgery. It can be used to study a wide range of topics including the economic burden of pediatric conditions, access to services, quality of care and patient safety, and the impact of health policy changes. Discharge weights are provided for calculating national estimates.

The KID target universe includes pediatric discharges from community, non-rehabilitation hospitals in the United States.¹ Pediatric discharges are defined as all discharges where a patient was 20 years or less at admission. Discharges with missing, invalid, or inconsistent ages are excluded. Pediatric discharges are identified as one of three types of records:

- Uncomplicated in-hospital births (HOSPBIRTH = 1 and UNCBIRTH = 1)
- Complicated in-hospital births (HOSPBIRTH = 1 and UNCBIRTH = 0)
- All other pediatric cases (HOSPBIRTH = 0).

In-hospital births (HOSPBIRTH = 1) are identified by any principal or secondary diagnosis code in the range of V3000 to V3901 with the last two digits of "00" or "01" and the patient is not transferred from another acute care hospital or health care facility. Uncomplicated births (UNCBIRTH = 1) have a Diagnosis Related Group (DRG) indicating "Normal Newborn" (391 prior to 2009, or 796 beginning in 2009).

The KID includes a sample of pediatric discharges from all HCUP hospitals in the sampling frame – the State Inpatient Databases (SID) that agreed to participate in the KID. For sampling, pediatric discharges are stratified by uncomplicated in-hospital birth, complicated in-hospital birth, and all other pediatric cases. To further ensure an accurate representation of each

¹ Community hospitals, as defined by the American Hospital Association (AHA), include "all non-Federal, short-term, general, and other specialty hospitals, excluding hospital units of institutions." Included among community hospitals are specialty hospitals such as obstetrics-gynecology, ear-nose-throat, short-term rehabilitation, orthopedic, and pediatric institutions. Also included are public hospitals and academic medical centers. Starting in 2005, the AHA included long term acute care facilities in the definition of community hospitals. These facilities provide acute care services to patients who need long term hospitalization (stays of more than 25 days). Excluded from the KID are short-term rehabilitation hospitals (beginning with 2000 data), long-term non-acute care hospitals, psychiatric hospitals, and alcoholism/chemical dependency treatment facilities.

hospital's pediatric case-mix, the discharges are sorted by hospital, DRG, and a random number within each DRG. Systematic random sampling is used to select 10% of uncomplicated in-hospital births and 80% of other pediatric cases from each frame hospital.

To obtain national estimates, discharge weights are developed using the American Hospital Association (AHA) universe of community, non-rehabilitation hospitals as the standard. For the weights, hospitals are post-stratified on six characteristics contained in the AHA hospital files—ownership/control, bed size, teaching status, rural/urban location, and U.S. region, with the addition of a stratum for freestanding children's hospitals. To create weights, if there were fewer than two frame hospitals, 30 uncomplicated births, 30 complicated births, and 30 non-birth pediatric discharges sampled in a stratum, that stratum is combined with an "adjacent" stratum containing hospitals with similar characteristics. Discharge weights are created by stratum in proportion to the number of AHA newborns for newborn discharges and in proportion to the total number of (non-newborn) AHA discharges for non-newborn discharges.

Detailed information on the design of the KID prior to 2006 is available in the year-specific special reports on [Design of the Kids' Inpatient Database](http://hcup-us.ahrq.gov/db/nation/kid/kidrelatedreports.jsp) found on the HCUP-US Web site (<http://hcup-us.ahrq.gov/db/nation/kid/kidrelatedreports.jsp>). Starting with the 2006 KID, the information on the design of the KID was incorporated into this report, which describes the KID sample and weights, summarizes the contents of the KID, and discusses data analysis issues. This document highlights cumulative information for all previous KID releases to provide a longitudinal view of the database. Over time, we have enhanced the nationwide representation of the sample by incorporating data from additional HCUP State Partners.

KID data sets are currently available for multiple years. See [Table 3](#) of [Appendix I](#) for a summary of KID releases. Each release of the KID includes:

- Data in fixed-width ASCII format on DVD or CD-ROM (prior to 2009).
- 2 million to 3 million pediatric inpatient records per year.
- 2,500 to 4,100 hospitals per year (all SID hospitals with pediatric discharges).
- Discharge-level weights to calculate national estimates.
- Hospital File with hospital characteristics.
- KID Documentation and tools – including file specifications, programming source code for loading ASCII data into SAS and SPSS, and value labels. Beginning in 2006, code is also provided for loading the KID ASCII file into Stata.

KID Data Sources, Hospitals, and Inpatient Stays

[Table 2](#) in [Appendix I](#) contains a summary of the data sources, number of hospitals, and number of inpatient stays in each KID database. It also lists the differences in types of hospitals and age inclusion for pediatric cases.

State-Specific Restrictions

Some HCUP Partners that contributed data to the KID imposed restrictions on the release of certain data elements or on the number and types of hospitals that could be included in the database. Because of confidentiality laws, some data sources were prohibited from providing HCUP with discharge records that indicated specific medical conditions, such as HIV/AIDS or behavioral health. Detailed information on these State-specific restrictions is available in [Appendix II](#).

Contents of KID

The KID is distributed as fixed-width ASCII formatted data files compressed with SecureZIP® from PKWARE. **Beginning with the 2012 KID, the files are encrypted.** Previously it was distributed on CD-ROMs, but beginning with the 2009 KID, it is distributed on a single DVD. It includes the following files:

Inpatient Core File: The Core file contains pediatric discharges sampled from community, non-rehabilitation hospitals in participating HCUP States. The unit of observation is an *inpatient stay record*. The Core file contains data elements for linkage, patient demographics, clinical information, and payment information. Sample weights for the three types of records, uncomplicated in-hospital births, complicated in-hospital births, and all other pediatric cases, are calculated separately by stratum and are added to each discharge in the Core File, as appropriate, so that only one discharge weight data element (DISCWT) is needed. See [Table 1](#) of [Appendix III](#) for a list of data elements in the Inpatient Core File. This file is available in all years of the KID.

Hospital File: The hospital-level file contains one observation for each hospital included in the KID and contains variance estimation data elements, linkage data elements, and data elements that describe basic characteristics about hospitals. The unit of observation is the *hospital*. Prior to the 2012 KID, the HCUP hospital identifier (HOSPID) provided the linkage between the KID Inpatient Core files and the Hospital Weights file. ***Beginning with the 2012 KID, the KID hospital number (HOSP_KID) provides the linkage between the KID Inpatient Core files and the Hospital File. The HOSP_KID values are reassigned each year, so they cannot be used to link hospitals across years.*** See [Table 2](#) of [Appendix III](#) for a list of data elements in the Hospital File. This file is available in all years of the KID.

Disease Severity Measures File: This discharge-level file contains information from two different sets of disease severity measures (APR-DRGs and Comorbidity measures). Information from the severity file is to be used in conjunction with the inpatient Core file. The unit of observation is an *inpatient stay record*. The HCUP unique record identifier (RECNUM) provides the linkage between the Core file and the Disease Severity Measures files. See [Appendix III](#), [Table 3](#) for a list of data elements in the Severity Measures Files. This file is available beginning with the 2003 KID.

Diagnosis and Procedure Groups File: This discharge-level file contains data elements from AHRQ software tools designed to facilitate the use of the ICD-9-CM diagnostic and procedure information in the HCUP databases. The unit of observation is an *inpatient stay record*. The HCUP unique record identifier (RECNUM) provides the linkage between the Core file and the Diagnosis and Procedure Groups file. [Table 4](#) in [Appendix III](#) contains a list of data elements in the Diagnosis and Procedure Groups file. This file is available beginning with the 2006 KID.

On the [HCUP-US](http://www.hcup-us.ahrq.gov) Web site (<http://www.hcup-us.ahrq.gov>), KID purchasers can access complete file documentation, including data element notes, file layouts, summary statistics, and related technical reports. Similarly, purchasers can also download SAS, SPSS, and Stata load programs. Available online documentation and supporting files are detailed in [Appendix I](#), [Table 4](#).

KID Data Elements

The KID contains two types of data: inpatient stay core records and hospital information. [Appendix III](#) identifies the data elements in each KID file:

- [Table 1](#) for the Inpatient Core files (record = inpatient stay)
- [Table 2](#) for the Hospital Weights files (record = hospital)
- [Table 3](#) for the Disease Severity Measures files (record = inpatient stay). This file was added beginning with the 2003 KID.
- [Table 4](#) for the Diagnosis and Procedure Groups files (record = inpatient stay). This file was added beginning with the 2006 KID.

The tables in [Appendix III](#) are not complete documentation for the data. Please refer to the KID documentation located on the [HCUP-US](#) Web site (<http://hcup-us.ahrq.gov>) for comprehensive information about data elements and the files.

Getting Started

In order to load and analyze the KID data on a computer, you will need the following:

- A DVD drive
- A hard drive with at least 10 gigabytes of space available
- SAS, SPSS, Stata or similar analysis software.

Copying and Decompressing the ASCII Files

To copy and decompress the data from the DVD, follow these steps:

- 1) Create a directory for the KID on your hard drive.
- 2) Unzip each ASCII file from the DVD, saving it into the new directory using a third-party zip utility such as ZIP Reader, WinZip, or 7-Zip. Beginning with the 2012 KID, you will be prompted to enter the encryption password (sent separately by email) to decrypt each file.

NOTE: Attempts to unzip encrypted files using Windows will produce an error message similar to the following: "Compressed (Zipped) Folders Error. The password you typed is not correct. Try typing it again." The solution is to use a third-party zip utility such as ZIP Reader, 7-Zip, or WinZip rather than the built-in Windows Explorer function to open the archive.

ZIP Reader may be downloaded for free at <http://comm.pkware.com/download-zipreader-business.html>.

7-Zip can be downloaded for free at <http://www.7-zip.org/>.

Evaluation versions of WinZip may be downloaded from the WinZip Web site at www.winzip.com.

Downloading and Running the Load Programs

Programs to load the data into SAS, SPSS, or Stata, are available on the HCUP User Support Web site (HCUP-US). To download and run the load programs, follow these steps:

- 1) Go to the [KID Database Documentation](http://www.hcup-us.ahrq.gov/db/nation/kid/kidbdbdocumentation.jsp) page on HCUP-US at <http://www.hcup-us.ahrq.gov/db/nation/kid/kidbdbdocumentation.jsp>.

- 2) Go to the “Load Programs” section on this page.
- 3) Click on “SAS Load Programs”, “SPSS Load Programs”, or “STATA Load Programs” to go to the corresponding Load Programs page.
- 4) Select and download the load programs you need. The load programs are specific to the data year. For example, the load program for the 2012 KID Core file is linked to “Core File” under “2012 KID”. Save the load programs into the same directory as the KID ASCII files on your computer.
- 5) Edit and run the load programs as appropriate for your environment to load and save the analysis files. For example, add directory paths for the input and output files if needed.

KID Documentation

KID documentation files on the [HCUP-US](http://hcup-us.ahrq.gov) Web site (<http://hcup-us.ahrq.gov>) provide important resources for the user. Refer to these resources to understand the structure and content of the KID and to aid in using the database.

- To locate the KID documentation on [HCUP-US](http://www.hcup-us.ahrq.gov), choose “Databases” from the home page (<http://www.hcup-us.ahrq.gov>). The section labeled “KIDS’ Inpatient Database (KID) is specific to the KID.

[Table 4](#) in [Appendix I](#) details both the KID related reports and the comprehensive KID documentation available on HCUP-US.

HCUP On-Line Tutorials²

For additional assistance, AHRQ has created the HCUP Online Tutorial Series, a series of free, interactive courses which provide training on technical methods for conducting research with HCUP data. Topics include an [HCUP Overview Course](#) and these tutorials:

The [Load and Check HCUP Data](#) tutorial provides instructions on how to unzip (decompress) HCUP data, save it on your computer, and load the data into a standard statistical software package. This tutorial also describes how to verify that the data have loaded correctly.

The [HCUP Sampling Design](#) tutorial is designed to help users learn how to account for sample design in their work with HCUP national (nationwide) databases.

The [Producing National HCUP Estimates](#) tutorial is designed to help users understand how the three national (nationwide) databases – the National Inpatient Sample (NIS), Nationwide Emergency Department Sample (NEDS), and KID – can be used to produce national and regional estimates.

The [Calculating Standard Errors](#) tutorial shows how to accurately determine the precision of the estimates produced from the HCUP nationwide databases. Users will learn two methods for calculating standard errors for estimates produced from the HCUP national (nationwide) databases.

² As of July, 2014, the HCUP Online Tutorials had not yet been updated for the 2012 KID data element changes. However, the same statistical techniques should be used to calculate standard errors and confidence intervals. There is one change in example programs: HOSPID (the encrypted hospital identifier) should be replaced by HOSP_KID.

The [HCUP Multi-year Analysis](#) tutorial presents solutions that may be necessary when conducting analyses that span multiple years of HCUP data.

New tutorials are added periodically and existing tutorials are updated when necessary. The [Online Tutorial Series](#) is located on the HCUP-US Web site at http://hcup-us.ahrq.gov/tech_assist/tutorials.jsp.

HOW TO USE THE KID FOR DATA ANALYSIS

This section provides a brief synopsis of special considerations when using the KID. For more details, refer to the comprehensive documentation on the [HCUP-US](#) Web site (<http://hcup-us.ahrq.gov>).

- If anyone other than the original purchaser uses the KID data, be sure to have them read and sign a Data Use Agreement, after viewing the on-line Data Use Agreement Training Tool available on the [HCUP-US](#) Web site (<http://www.hcup-us.ahrq.gov>). A copy of the signed Data Use Agreements must be sent to AHRQ. See page 2 for the mailing address.
- The KID contains discharge-level records, not patient-level records. This means that individual patients who are hospitalized multiple times in one year may be present in the KID multiple times. There is no uniform patient identifier available that allows a patient-level analysis with the KID. This will be especially important to remember for certain conditions for which patients may be hospitalized multiple times in a single year.

Calculating National Estimates

- **In order to produce national estimates, you MUST use discharge weights.** Use the discharge weight (DISCWT³) to project discharges in the KID Core files to the discharges from all U.S. community, non-rehabilitation hospitals. (For trends analysis using 1997 KID data, see the next section of this report regarding trend weights.)
- Proper statistical techniques must be used to calculate standard errors and confidence intervals when using the KID. For detailed instructions, refer to the special report [Calculating Kids' Inpatient Database \(KID\) Variances](#),⁴ available on www.hcup-us.ahrq.gov.

The KID Comparison Reports (available on www.hcup-us.ahrq.gov) assess the accuracy of KID estimates. KID Comparison reports are available for 1997 and 2003. These reports show that the KID's large sample size enables reliable national estimates.

- When creating national estimates, it is a good idea to check your estimates against other data sources, if available.

³ Prior to 2000, the discharge weight was named DISCWT_U. For 2000 only, use DISCWT to create national estimates for all analyses except those that involve total charges; and use DISCWTCHARGE to create national estimates of total charges.

⁴ As of July, 2014, this report had not yet been updated for the 2012 KID data element changes. However, the same statistical techniques should be used to calculate standard errors and confidence intervals. There is one change in example programs: HOSPID (the encrypted hospital identifier) should be replaced by HOSP_KID.

- To ensure that you are using the weights appropriately and calculating estimates and variances accurately, check your estimates against [HCUPnet](http://www.hcupnet.ahrq.gov), the free online query system (<http://www.hcupnet.ahrq.gov>). HCUPnet is a Web-based query tool for identifying, tracking, analyzing, and comparing statistics on hospitals at the national, regional, and State level. HCUPnet offers easy access to national statistics and trends and selected State statistics about hospital stays. This tool provides step-by-step guidance, helping researchers to quickly obtain the statistics they need. HCUPnet generates statistics using the NIS, KID, and SID for those States that have agreed to participate. In addition, HCUPnet provides Quick Statistics – ready-to-use tables on commonly requested information – as well as national statistics based on the AHRQ Quality Indicators.

Studying Trends

- When studying trends over time using the KID, be aware that the sampling frame for the KID changes over time (i.e., more States have been added). Estimates from earlier years of the KID may be subject to more sampling bias than later years of the KID.
- Short-term rehabilitation hospitals are included in the 1997 KID, but are excluded from later years of the KID. Patients treated in short-term rehabilitation hospitals tend to have lower mortality rates and longer lengths of stay than patients in other community hospitals. The elimination of rehabilitation hospitals may affect trends but the effect is likely small since only about 3% of community hospitals are short-term rehabilitation hospitals and not all State data sources included short term rehabilitation hospitals.
- In order to facilitate analysis of trends including the 1997 KID, an alternate set of KID discharge weights for the 1997 HCUP KID were developed. These trends weights were calculated in the same way as the weights for the 2000 and later years of the KID. (Trend analyses for 2000 and later data do not need the KID trends weights.) The report, [Using the Kids' Inpatient Database \(KID\) to Estimate Trends](#), includes details regarding the [KID trends weights](#) and other recommendations for trends analysis. [The KID trends report](#) is available on the HCUP-US Web site at http://www.hcup-us.ahrq.gov/reports/methods/2007_02.pdf, and the [KID trends weights](#) are available on the HCUP-US Web site at <https://www.hcup-us.ahrq.gov/db/nation/kid/kidtrends.jsp>.

Choosing Data Elements for Analysis

- For all data elements you plan to use in your analysis, first perform descriptive statistics and examine the range of values, including number of missing cases. Summary statistics for the entire KID are provided on the [Summary Statistics](#) page of the HCUP-US Web site (<http://hcup-us.ahrq.gov/db/nation/kid/kidsummarystats.jsp>). Performing descriptive statistics by hospital can be helpful in detecting hospital-specific data anomalies.
- Not all data elements in the KID are provided by each hospital. These data elements are provided on the KID because they can be valuable for research purposes but they should be used cautiously. For example, RACE is missing for some hospitals; thus, national estimates using RACE should be interpreted and reported with caveats.
- Differences exist across the State data sources in the collection of information that could not be accounted for during HCUP processing to make the data uniform. For example, the most reliable way to identify ED admissions in the HCUP databases is to use the data

element HCUP_ED, which considers all possible evidence of ED services. Unfortunately, this information is not always complete.

Hospital-Level Data Elements

- The 2012 Kids' Inpatient Database (KID) incorporates changes to enhance confidentiality—the 2012 KID eliminates State identifiers, unencrypted hospital identifiers (the AHA identifier), and other data elements that are not uniformly available across States. This means that you will not be able to link the KID to outside data sources that require hospital-specific identifiers. However, there are hospital-level data elements for all hospitals in the KID that allow you to study certain hospital characteristics including ownership/control, teaching status, rural/urban location, bed size, the nine census divisions (as well as the four regions), and a data element that allows you to identify whether a hospital is a freestanding children's hospital (KID_STRATUM=9999 or 9998). A detailed description of the data elements is available on [HCUP-US](#).

ICD-9-CM Diagnosis and Procedure Codes

- ICD-9-CM diagnosis and procedure codes provide valuable insights into the reasons for hospitalization and what procedures patients receive, but these codes need to be carefully used and interpreted. ICD-9-CM codes change every October as new codes are introduced and some codes are retired. See the Conversion Table at <http://www.cdc.gov/nchs/datawh/ftp/ftpicd9/ftpicd9.htm> which shows ICD-9-CM code changes over time. **It is critical to check all ICD-9-CM codes used for analysis to ensure the codes are in effect during the time period studied.**
- Although the KID contains up to 25 diagnoses (15 prior to the 2009 KID) and 15 procedures, the number of diagnoses and procedures varies across hospitals. Some hospitals provide as many as 31 diagnoses and procedures or more, while other hospitals provide as few as 9 diagnoses and 6 procedures. Because very few cases have more than 25 diagnoses or 15 procedures, the diagnosis and procedure vectors were truncated to save space in the KID data files. Two data elements are provided which tell you exactly how many diagnoses and procedures were on the original records (NDX and NPR). See the notes on [diagnoses](http://www.hcup-us.ahrq.gov/db/vars/dxn/kidnote.jsp) (<http://www.hcup-us.ahrq.gov/db/vars/dxn/kidnote.jsp>) and [procedures](http://www.hcup-us.ahrq.gov/db/vars/_prn/kidnote.jsp) (http://www.hcup-us.ahrq.gov/db/vars/_prn/kidnote.jsp) on the HCUP-US Web site.
- The collection and reporting of external cause of injury (E codes) varies across hospitals depending on the presence of State laws or mandates for the collection of E codes. Some States do not require hospitals to report E codes in the range E870-E879 - "misadventures to patients during surgical and medical care" - which means that these occurrences will be underreported.

Missing Values

Missing data values can compromise the quality of estimates. If the outcome for discharges with missing values is different from the outcome for discharges with valid values, then sample estimates for that outcome will be biased and inaccurately represent the discharge population. For example, race is missing on about 8% of discharges in the 2012 KID because some hospitals and HCUP State Partners do not supply it. (The percentage of missing race values was higher in previous years.) Therefore race-specific estimates may be biased. This is especially true for estimates of discharge totals by race.

There are several techniques available to help overcome this bias.⁵ Descriptions of such data preparation and adjustment are outside the scope of this report; however, it is recommended that researchers evaluate and adjust for missing data, if necessary.

Variance Calculations

It may be important for researchers to calculate a measure of precision for some estimates based on the KID sample data. Variance estimates must take into account both the sampling design and the form of the statistic. If hospitals inside the frame are similar to hospitals outside the frame, the sample hospitals can be treated as if they were randomly selected from the entire universe of hospitals within each stratum. Discharges were randomly selected from within each hospital. Standard formulas for stratified, two-stage cluster samples without replacement may be used to calculate statistics and their variances in most applications. **To accurately calculate variances from the KID, you must use appropriate statistical software and techniques.** For details, see the special report, [Calculating Kids' Inpatient Database \(KID\) Variances](http://www.hcup-us.ahrq.gov/db/nation/kid/kidrelatedreports.jsp).⁶ This report is available on the HCUP-US Web site at <http://www.hcup-us.ahrq.gov/db/nation/kid/kidrelatedreports.jsp>.

A multitude of statistics can be estimated from the KID data. Several computer programs that calculate statistics and their variances from sample survey data are listed in the section below. Some of these programs use general methods of variance calculations (e.g., the jackknife and balanced half-sample replications) that take into account the sampling design. However, it may be desirable to calculate variances using formulas specifically developed for some statistics.

These variance calculations are based on finite-sample theory, which is an appropriate method for obtaining cross-sectional, nationwide estimates of outcomes. According to finite-sample theory, the intent of the estimation process is to obtain estimates that are precise representations of the nationwide population at a specific point in time. In the context of the KID, any estimates that attempt to accurately describe characteristics (such as expenditure and utilization patterns or hospital market factors) and interrelationships among characteristics of hospitals and discharges during a specific year should be governed by finite-sample theory.

Alternatively, in the study of hypothetical population outcomes not limited to a specific point in time, the concept of a “superpopulation” may be useful. Analysts may be less interested in specific characteristics from the finite population (and time period) from which the *sample* was drawn than they are in hypothetical characteristics of a conceptual superpopulation from which any particular finite *population* in a given year might have been drawn. According to this superpopulation model, the nationwide population in a given year is only a snapshot in time of the possible interrelationships among hospital, market, and discharge characteristics. In a given year, all possible interactions between such characteristics may not have been observed, but analysts may wish to predict or simulate interrelationships that may occur in the future.

⁵ See, for example, van Buuren, S. (2012). *Flexible Imputation of Missing Data*. CRC Press, Boca Raton, FL.

⁶ As of July, 2014, this report had not yet been updated for the 2012 KID data element changes. However, the same statistical techniques should be used to calculate standard errors and confidence intervals. There is one change in example programs: HOSPID (the encrypted hospital identifier) should be replaced by HOSP_KID.

Under the finite-population model, the variances of estimates approach zero as the sampling fraction approaches one. This is the case because the population is defined at that point in time, and because the estimate is for a characteristic as it existed when sampled. This is in contrast to the superpopulation model, which adopts a stochastic viewpoint rather than a deterministic viewpoint. That is, the nationwide population in a particular year is viewed as a random sample of some underlying superpopulation over time. Different methods are used for calculating variances under the two sample theories. The choice of an appropriate method for calculating variances for nationwide estimates depends on the type of measure and the intent of the estimation process.

Computer Software for Variance Calculations

The discharge weights would be used to weight the sample data in estimating population statistics. In most cases, computer programs are readily available to perform these calculations. Several statistical programming packages allow weighted analyses.⁷ For example, nearly all SAS procedures incorporate weights. In addition, several statistical analysis programs have been developed to specifically calculate statistics and their standard errors from survey data. Version eight or later of SAS contains procedures (PROC SURVEYMEANS and PROC SURVEYREG) for calculating statistics based on specific sampling designs. STATA and SUDAAN are two other common statistical software packages that perform calculations for numerous statistics arising from the stratified, single-stage cluster sampling design. Examples of the use of SAS, SUDAAN, and STATA to calculate KID variances are presented in the special report: [Calculating Kids' Inpatient Database \(KID\) Variances](#). This report is available on the HCUP-US Web site at <http://www.hcup-us.ahrq.gov/db/nation/kid/kidrelatedreports.jsp>. For an excellent review of programs to calculate statistics from survey data, visit the following Web site: <http://www.hcp.med.harvard.edu/statistics/survey-soft/>.

The KID database includes a Hospital file with data elements required to calculate finite population statistics. The file includes hospital identifiers (Primary Sampling Units or PSUs), stratification data elements, and stratum-specific totals for the numbers of discharges and hospitals so that finite-population corrections can be applied to variance estimates.

In addition to these subroutines, standard errors can be estimated by validation and cross-validation techniques. Given that a very large number of observations will be available for most analyses, it may be feasible to set aside a part of the data for validation purposes. Standard errors and confidence intervals can then be calculated from the validation data.

If the analytical file is too small to set aside a large validation sample, cross-validation techniques may be used. For example, tenfold cross-validation would split the data into ten equal-sized subsets. The estimation would take place in ten iterations. In each iteration, the outcome of interest is predicted for one-tenth of the observations by an estimate based on a model fit to the other nine-tenths of the observations. Unbiased estimates of error variance are then obtained by comparing the actual values to the predicted values obtained in this manner.

Finally, it should be noted that a large array of hospital-level data elements are available for the entire universe of hospitals, including those outside the sampling frame. For instance, the data elements from the AHA surveys and from the Medicare Cost Reports are available for nearly all

⁷ Carlson BL, Johnson AE, Cohen SB. "An Evaluation of the Use of Personal Computers for Variance Estimation with Complex Survey Data." *Journal of Official Statistics*, vol. 9, no. 4, 1993: 795-814.

hospitals in the U.S, although hospital identifiers are suppressed in the KID for a number of States. For these States it will not be possible to link to outside hospital-level data sources. To the extent that hospital-level outcomes correlate with these data elements, they may be used to sharpen regional and nationwide estimates.

SAMPLING OF DISCHARGES

Sampling of Discharges Included in the KID

Unlike the HCUP Nationwide Inpatient Sample (NIS) prior to 2012, the KID has never involved sampling hospitals. Instead, the KID includes a sample of pediatric discharges from all hospitals in the sampling frame⁸. For the sampling, pediatric discharges in all participating States are stratified by uncomplicated in-hospital birth, complicated in-hospital birth, and all other pediatric cases. To further ensure an accurate representation of each hospital's pediatric case-mix, the discharges are sorted by State, hospital, DRG, and a random number within each DRG. Systematic random sampling is used to select 10% of uncomplicated in-hospital births and 80% of complicated in-hospital births and other pediatric cases from each frame hospital.

To obtain national estimates, discharge weights are developed using the AHA universe as the standard. For the weights, hospitals are post-stratified on six characteristics contained in the AHA hospital files—ownership/control, bed size, teaching status, rural/urban location, and U.S. region, with the addition of a stratum for freestanding children's hospitals. If there were fewer than two frame hospitals, 30 uncomplicated births, 30 complicated births, and 30 non-birth pediatric discharges sampled in a stratum, that stratum is combined with an "adjacent" stratum containing hospitals with similar characteristics. Discharge weights are created by stratum in proportion to the total number of AHA newborns for in-hospital births and in proportion to the total number of AHA non-newborn admissions for non-birth pediatric discharges.

The KID Hospital Universe

The hospital universe is defined as all hospitals located in the U.S. that were open during any part of the calendar year and that were designated as community hospitals in the AHA Annual Survey Database. The AHA defines community hospitals as follows: "All non-Federal, short-term, general, and other specialty hospitals, excluding hospital units of institutions." Starting in 2005, the AHA included long term acute care facilities in the definition of community hospitals. These facilities provide acute care services to patients who need long term hospitalization (more than 25 days stays). Consequently, Veterans Hospitals and other Federal facilities (Department of Defense and Indian Health Service) are excluded. Beginning with the 2000 KID, short-term rehabilitation hospitals were excluded from the universe, because the type of care provided and the characteristics of the discharges from these facilities were markedly different from other short-term hospitals. (The 1997 KID includes short-term rehabilitation hospitals. The KID Trend Weights, described earlier in this report, remove these hospitals and adjust for other design changes in the 2000 KID.) [Table 2](#) in [Appendix I](#) displays the number of hospitals in the universe for each year, based on the corresponding AHA Annual Survey Database.

⁸ As of 2012, the sampling strategy for National Inpatient Sample (NIS) was redesigned as a sample of discharges from all HCUP-participating hospitals. For more information on the new design for the NIS, see the [NIS Overview](#), available on HCUP-US Web site at <http://www.hcup-us.ahrq.gov/nisoverview.jsp>.

For more information on how hospitals in the data set were mapped to hospitals as defined by the AHA, refer to the special report, [HCUP Hospital Identifiers](#). For a list of all data sources, refer to [Table 1](#) in [Appendix I](#). Detailed information on the design of the KID prior to 2006 is available in the year-specific special reports on [Design of the Kids' Inpatient Database](#) found on the HCUP-US Web site at <http://www.hcup-us.ahrq.gov/db/nation/kid/kidrelatedreports.jsp>. Starting with the 2006 KID, the design information was incorporated into this report.

Hospital Merges, Splits, and Closures

All U.S. hospital entities that were designated community hospitals in the AHA hospital file, except short-term rehabilitation hospitals, were included in the hospital universe. Therefore, when two or more community hospitals merged to create a new community hospital, the original hospitals and the newly-formed hospital were all considered separate hospital entities in the universe during the year they merged. Similarly, if a community hospital split, the original hospital and all newly-created community hospitals were treated as separate entities in the universe during the year this occurred. Finally, community hospitals that closed during a given year were included in the hospital universe, as long as they were in operation during some part of the calendar year.

Stratification Data Elements

For the purpose of calculating discharge weights, we post-stratified hospitals on six characteristics contained in the AHA hospital files—ownership/control, bed size, teaching status, rural/urban location, and U.S. region, with the addition of a stratum for freestanding children's hospitals. The definitions of some of the strata were revised beginning with the 2000 KID. (A description of the strata used for the 1997 KID can be found in the [Kids' Inpatient Database \(KID\) Design Report, 1997](#). This report is available on the HCUP-US Web site at <http://www.hcup-us.ahrq.gov/db/nation/kid/kidrelatedreports.jsp>.)

Beginning with the 2000 KID, the stratification data elements were defined as follows:

- *Geographic Region – Northeast, Midwest, West, and South.* This is an important stratification data element because practice patterns have been shown to vary substantially by region. For example, lengths of stay tend to be longer in East Coast hospitals than in West Coast hospitals. [Figure 1](#) in [Appendix I](#) shows the KID States by region.
- *Control – government non-Federal (public), private not-for-profit (voluntary), and private investor-owned (proprietary).* These types of hospitals tend to have different missions and different responses to government regulations and policies. When there were enough hospitals of each type to allow it, hospitals were stratified as public, voluntary, and proprietary. This stratification was used for Southern rural, Southern urban non-teaching, and Western urban non-teaching hospitals. For smaller strata – the Midwestern rural and Western rural hospitals – a collapsed stratification of public versus private was used, with the voluntary and proprietary hospitals combined to form a single “private” category. For all other combinations of region, location, and teaching status, no stratification based on control was advisable, given the number of hospitals in these cells.
- *Location – urban or rural.* Government payment policies for hospital services often differ according to this designation. Also, rural hospitals are generally smaller and offer fewer services than urban hospitals. We categorized hospitals with a CBSA type of *Metropolitan* or

Division as urban, while we designated hospitals with a CBSA type of *Micropolitan* or *Rural* as rural.

- *Teaching Status – teaching or non-teaching.* The missions of teaching hospitals differ from non-teaching hospitals. In addition, financial considerations differ between these two hospital groups. Currently, the Medicare DRG payments are uniformly higher to teaching hospitals. Prior to 2006, the teaching status of hospitals identified as children’s hospitals by the Children’s Hospital Association (CHA, formerly National Association of Children’s Hospitals and Related Institutions, NACHRI) was based on an indicator provided by CHA. For 2006, the CHA teaching status indicator was not available, so teaching status was determined using only information from the AHA Annual Survey Database for all hospitals. For 2009, both CHA and AHA information were used to define the teaching status of children’s hospitals. For 2012, the CHA teaching status indicator was not available, so teaching status was determined using only information from the AHA Annual Survey Database for all hospitals.

In the 1997 KID, we considered other hospitals to be teaching hospitals if they had any residents or interns and met one of the following two criteria:

- Residency training approval by the Accreditation Council for Graduate Medical Education (ACGME)
- Membership in the Council of Teaching Hospitals (COTH).

Beginning with the KID, we considered other hospitals to be teaching hospitals if they met any one of the following three criteria: (See [Appendix IV](#) for details.)

- Residency training approval by the Accreditation Council for Graduate Medical Education (ACGME)
 - Membership in the Council of Teaching Hospitals (COTH)
 - A ratio of full-time equivalent interns and residents to beds of .25 or greater.
- *Bed Size – small, medium, and large.* Bed size categories are based on hospital beds and are specific to the hospital's region, location, and teaching status, as illustrated in [Table 5 of Appendix I](#). The bed size cutoff points were chosen so that approximately one-third of the hospitals in a given region, location, and teaching status combination would fall within each bed size category (small, medium, or large). Different cutoff points for rural, urban non-teaching, and urban teaching hospitals were used because hospitals in those categories tend to be small, medium, and large, respectively. For example, a medium-sized teaching hospital would be considered a rather large rural hospital. Further, the size distribution is different among regions for each of the urban/teaching categories. For example, teaching hospitals tend to be smaller in the West than they are in the South. Using differing cutoff points in this manner avoids strata containing small numbers of hospitals.

Rural hospitals were not split according to teaching status, because rural teaching hospitals were rare. For example, rural teaching hospitals generally comprise about 2% or less of the total hospital universe. The bed size categories were defined within location and teaching status because they would otherwise have been redundant. Rural hospitals tend to be small; urban non-teaching hospitals tend to be medium-sized; and urban teaching hospitals tend to be large. Yet it was important to recognize gradations of size within these types of hospitals.

For example, in serving rural discharges, the role of "large" rural hospitals (particularly rural referral centers) often differs from the role of "small" rural hospitals.

- *Hospital Type – freestanding children’s or other hospital.* Children’s hospitals restrict admissions to children, while other hospitals admit both adults and children. There may be significant differences in practice patterns, casemix, severity of illness, and available services between children’s hospitals and other hospitals. Freestanding children’s hospitals are assigned to a separate stratum and can be identified using KID_STRATUM=9999 (or 9998 in 2003). Children’s units in general hospitals were not stratified as children’s hospitals. Data from the Children’s Hospital Association (CHA) were used to help verify the AHA list of children’s hospitals.

Hospital Sampling Frame

The *universe* of hospitals was established as all community hospitals located in the U.S. with the exception, beginning in 2000, of short-term rehabilitation hospitals. However, some hospitals do not supply data to HCUP. Therefore, we constructed the KID *sampling frame* from the subset of universe hospitals that released their discharge data to AHRQ for research use. The number of State Partners and hospitals contributing data to the KID has expanded over the years, as shown in [Table 2](#) of [Appendix I](#).

The list of the entire frame of hospitals was composed of all AHA community, non-rehabilitation hospitals in each of the frame States *that could be matched to the discharge data provided to HCUP*. If an AHA hospital could not be matched to the discharge data provided by the data source, it was eliminated from the sampling frame (but not from the target universe).

[Table 6](#) of [Appendix I](#) shows the number of AHA, HCUP SID, and KID hospitals by Region and by Census Division. In most cases, the difference between the universe and the frame represents the difference between the number of community, non-rehabilitation hospitals in the 2012 AHA Annual Survey Database and the number of hospitals with children’s discharges that were supplied to HCUP that could be matched to the AHA data.

Beginning with the 2000 KID, pediatric discharges were defined as having an age at admission of 20 or less. This differs from the 1997 KID, which included discharges with an admission age of 18 or less. Discharges with missing, invalid, or inconsistent ages were excluded.

Hospital Sample Design

Design Considerations

The overall design objective was to select a sample of pediatric discharges that accurately represents the target universe of U.S. community, non-rehabilitation hospitals. Moreover, this sample was to be geographically dispersed, yet drawn exclusively from hospitals in States that participate in HCUP.

It should be possible, for example, to estimate DRG-specific average lengths of stay across all U.S. hospitals using weighted average lengths of stay, based on averages or regression coefficients calculated from the KID. Ideally, relationships among outcomes and their correlates estimated from the KID should accurately represent all U.S. hospitals. It is advisable to verify

your estimates against other data sources, especially for specific patient populations (e.g. organ transplant recipients).

In order to sample and project births up to the number of births reported by the AHA, which reports in-hospital births, the KID development team identified all in-hospital births in the KID data. We further separated the in-hospital births in HCUP data into uncomplicated births and complicated births. We sampled uncomplicated births at a lower rate because they have little variation in their outcomes.

To determine the best way to identify in-hospital births, we ran cross-tabulations of different combinations of data elements on all cases that had any of the following possible birth indicators: age of zero days (AGEDAY=0), neonatal diagnosis (NEOMAT>=2), neonatal Major Diagnostic Category (MDC 15), or admission type of birth (ATYPE=4).⁹ Based on reviews of the cross-tabulations, the MDC 15 DRG definitions, and ICD-9-CM birth diagnosis codes, the following screen was devised for births: an in-hospital birth diagnosis code (any diagnosis code in the range V3000 - V3901 with a fourth digit of zero, indicating born in the hospital, and a fifth digit of zero or one, indicating delivered without mention of cesarean delivery, or delivered by cesarean delivery), without an admission source of another hospital or health facility (ASOURCE not equal to 2 or 3).

We classified neonates transferred from other facilities as pediatric non-births because they are not included in births reported by the AHA. An age of zero days was not a reliable in-hospital birth indicator because neonates transferred from another hospital or born before admission to the hospital could also have an age of zero days. There were also some cases with birth diagnoses, but with ages of a few days. Because the HCUP data are already edited for neonatal diagnoses inconsistent with age, we did not include any age criteria in the in-hospital birth screen.

Uncomplicated, in-hospital births are identified as cases that meet the above screen and have a Diagnosis Related Group (DRG) indicating "Normal Newborn" (391 prior to 2009, or 796 beginning in 2009). In the KID, a small percentage of the cases with a DRG of "Normal Newborn" do not meet the in-hospital birth screen. These cases have diagnoses that imply a newborn, but do not specifically indicate an in-hospital birth. It is possible that some of these may have actually been born in the hospital but lacked the proper diagnosis code. Others may be readmissions or may have been born before admission to the hospital. Some of these cases have an admission type of newborn (ATYPE = 4).

Changes to Sampling and Weighting Strategy Beginning with the 2000 KID

We revised some of the hospital universe and strata definitions beginning with the 2000 KID. These changes included:

- Revising definitions of the strata data elements
- Excluding rehabilitation hospitals from the hospital universe
- Changing the calculation of hospital universe discharges for the weights.

⁹ We performed this analysis during the development of the original 1997 KID.

Sampling Procedure

The KID includes a sample of pediatric discharges from all hospitals in the sampling frame. For the sampling, we stratified the pediatric discharges by uncomplicated in-hospital birth, complicated in-hospital birth, and pediatric non-birth. To further ensure an accurate representation of each hospital's pediatric case-mix, we also sorted the discharges by State, hospital, DRG, and a random number within each DRG. We then used systematic random sampling to select 10% of "normal newborns" born in the hospital and 80% of other pediatric cases from each frame hospital.

It should be observed that the KID includes fewer than 100% of the pediatric discharges for each hospital in the database. Therefore, researchers will not be able to calculate hospital-specific outcomes with certainty.

SAMPLE WEIGHTS

To obtain national estimates, we developed discharge weights using the AHA universe as the standard. For the weights, hospitals are post-stratified on six characteristics contained in the AHA hospital files—ownership/control, bed size, teaching status, rural/urban location, and U.S. region, with the addition of a stratum for freestanding children's hospitals. We also stratified the KID discharges according to whether the discharge was an uncomplicated in-hospital birth, a complicated in-hospital birth, or a non-newborn pediatric discharge. If there were fewer than two frame hospitals, 30 uncomplicated births, 30 complicated births, and 30 non-birth pediatric discharges sampled in a stratum, we merged that stratum with an "adjacent" stratum containing hospitals with similar characteristics.

We used Children's Hospital Association (CHA) data to help verify the AHA list of children's hospitals in the target universe. Data analysts may find it useful to identify discharges from children's hospitals. Prior to 2012, children's hospitals within general hospitals were not stratified as children's hospitals, but they could be selected using the NACHTYPE data element in the KID. Beginning with 2012 data, NACHTYPE is no longer available, but discharges from freestanding children's hospitals are stratified as children's hospitals (KID_STRATUM=9999 or 9998).

Discharge Weights

The discharge weights usually are constant for all discharges of the same type (uncomplicated in-hospital birth, complicated in-hospital birth, and other pediatric discharge) within a stratum. The only exceptions are for strata with sample hospitals that, according to the AHA files, were open for the entire year but contributed less than their full year of data to the KID. For those hospitals, we *adjusted* the number of observed discharges by a factor of $4 \div Q$, where Q was the number of calendar quarters that the hospital contributed discharges to the KID. For example, when a sample hospital contributed only two quarters of discharge data to the KID, the *adjusted* number of discharges was double the observed number.

With that minor adjustment, each discharge weight is essentially equal to the number of AHA universe discharges that each sampled discharge represents in its stratum. This calculation was possible because the numbers of total discharges and births were available for every hospital in the universe from the AHA files.

Discharge weights to the universe were calculated by post-stratification. Hospitals were stratified on geographic region, urban/rural location, teaching status, bed size, control, and hospital type. In some instances, strata were collapsed for sample weight calculations. Within stratum k , for hospital i , each KID sample discharge's universe weight was calculated as:

$$W_{ik} = [T_k / (R_k * A_k)] * (4 \div Q_i)$$

In the birth strata (both complicated and uncomplicated):

- T_k is the total number of births reported in the AHA survey.
- A_k is the total number of adjusted births in the restricted sampling frame.
- In the uncomplicated birth strata, R_k is the frame sampling rate for uncomplicated in-hospital births calculated as the sum of the adjusted number of uncomplicated births sampled divided by the sum of the adjusted number of uncomplicated births in the restricted frame.
- In the complicated birth strata, R_k is the frame sampling rate for complicated in-hospital births.

In the non-newborn strata:

- T_k is the total number of non-newborns reported in the AHA survey.
- A_k is the total number of adjusted non-newborn discharges in the sampling frame.
- R_k is the frame sampling rate for non-newborns from all non-newborn discharges in the sampling frame.
- Q_i is the number of quarters of discharge data contributed by hospital i to the KID (usually $Q_i = 4$).
- T_k / A_k estimates the number of discharges in the population that is represented by each discharge in the sampling frame. R_k adjusts for the fact that we are taking a sample of the frame in each stratum.

Uncomplicated in-hospital births were sampled at a lower rate than other discharges because the variation in hospital outcomes for uncomplicated births is considerably less than that for other pediatric cases and because we expect research to focus much more on other pediatric patients. We sampled uncomplicated births at the nominal rate of 10% and sampled other pediatric discharges (complicated newborns and other pediatric cases) at the nominal rate of 80% from the discharges available in the (restricted) frame. To avoid rounding errors in the weights calculation, the actual sampling rate for a discharge type (uncomplicated in-hospital birth, complicated in-hospital birth, or non-birth pediatric discharge) in stratum k , R_k , was calculated as follows:

$$R_k = S_k / H_k$$

- S_k is the number of adjusted discharges **sampl**ed for the discharge type in stratum k .
- H_k is the number of adjusted discharges in the sampling frame for the discharge type in stratum k .

The AHA birth counts include both uncomplicated and complicated births. Therefore, the weights in the uncomplicated birth strata implicitly assume that the proportion of births that are uncomplicated in the frame is representative of the proportion of births that are uncomplicated in the population for each stratum. A similar assumption is made for complicated newborns.

Similarly, the non-birth AHA counts include all non-birth admissions, not just non-birth pediatric counts. Consequently, the weights in the non-birth strata implicitly assume that the proportion of non-birth discharges that are pediatric across the HCUP SID hospitals is the same as the proportion of non-birth admissions that are pediatric across the universe of AHA hospitals, in the aggregate within each hospital stratum.

Weight Data Elements

To produce nationwide estimates, use the discharge weights to project sampled discharges in the Core file to the discharges from all U.S. community, non-rehabilitation hospitals. Beginning with the 2003 KID, use DISCWT to calculate nationwide estimates for all analyses. For the 2000 KID, use DISCWT to create nationwide estimates for all analyses except those that involve total charges, and use DISCWTCHARGE to create nationwide estimates of total charges. For the 1997 KID, use DISCWT_U for all analyses. (For trends analysis using 1997 KID data, see the previous section of this report regarding "[Studying Trends](#).”)

THE FINAL KID SAMPLE

In [Appendix I](#), we present tables and figures that summarize the final KID sample.

[Table 2](#) summarizes information across all years of the KID, including the KID States, data sources, sample hospitals, and sample discharges.

[Table 7](#) shows the number of hospitals and discharges for children’s hospitals and other hospitals. For each hospital type, the table shows the number of:

- AHA universe hospitals and total admissions plus births
- Non-rehabilitation community hospitals in the SID and associated pediatric discharges
- Hospitals and pediatric discharges included in the KID.

[Table 8](#) displays the unweighted and weighted number of uncomplicated births, complicated births, and pediatric non-births by hospital type in the KID.

[Figure 2](#) displays the KID hospitals by geographic region. For each region, the chart presents:

- The number of hospitals in the AHA universe
- The number of SID community hospitals with pediatric discharges
- The number of hospitals in the KID (and the percentage of AHA universe hospitals).

Although pediatric discharges from hospitals in each region are selected for the KID, the comprehensiveness of the sampling frame varies by region, as shown in [Figure 2](#).

[Figure 3](#) summarizes the estimated U.S. population by geographic region on July 1, 2012. For each region, the figure reveals:

- The estimated U.S. population
- The estimated population of States in the KID
- The percentage of estimated U.S. population included in KID States.

This figure shows that the sampling frame for the KID includes states that comprise over 95 percent of the U.S. population.

Appendix I: Tables and Figures

Table 1. Data Sources for the 2012 KID

State	Data Organization
AK	Alaska State Hospital and Nursing Home Association
AR	Arkansas Department of Health
AZ	Arizona Department of Health Services
CA	Office of Statewide Health Planning & Development
CO	Colorado Hospital Association
CT	Connecticut Hospital Association
FL	Florida Agency for Health Care Administration
GA	Georgia Hospital Association
HI	Hawaii Health Information Corporation
IA	Iowa Hospital Association
IL	Illinois Department of Public Health
IN	Indiana Hospital Association
KS	Kansas Hospital Association
KY	Kentucky Cabinet for Health and Family Services
LA	Louisiana Department of Health and Hospitals
MA	Division of Health Care Finance and Policy
MD	Health Services Cost Review Commission
MI	Michigan Health & Hospital Association
MN	Minnesota Hospital Association
MO	Hospital Industry Data Institute
MT	MHA - An Association of Montana Health Care Providers
NC	North Carolina Department of Health and Human Services
ND	North Dakota (data provided by the Minnesota Hospital Association)

State	Data Organization
NE	Nebraska Hospital Association
NJ	New Jersey Department of Health
NM	New Mexico Department of Health
NV	Nevada Department of Health and Human Services
NY	New York State Department of Health
OH	Ohio Hospital Association
OK	Oklahoma State Department of Health
OR	Oregon Association of Hospitals and Health Systems
PA	Pennsylvania Health Care Cost Containment Council
RI	Rhode Island Department of Health
SC	South Carolina State Budget & Control Board
SD	South Dakota Association of Healthcare Organizations
TN	Tennessee Hospital Association
TX	Texas Department of State Health Services
UT	Utah Department of Health
VT	Vermont Association of Hospitals and Health Systems
VA	Virginia Health Information
WA	Washington State Department of Health
WI	Wisconsin Department of Health Services
WV	West Virginia Health Care Authority
WY	Wyoming Hospital Association

[Return to Introduction](#)

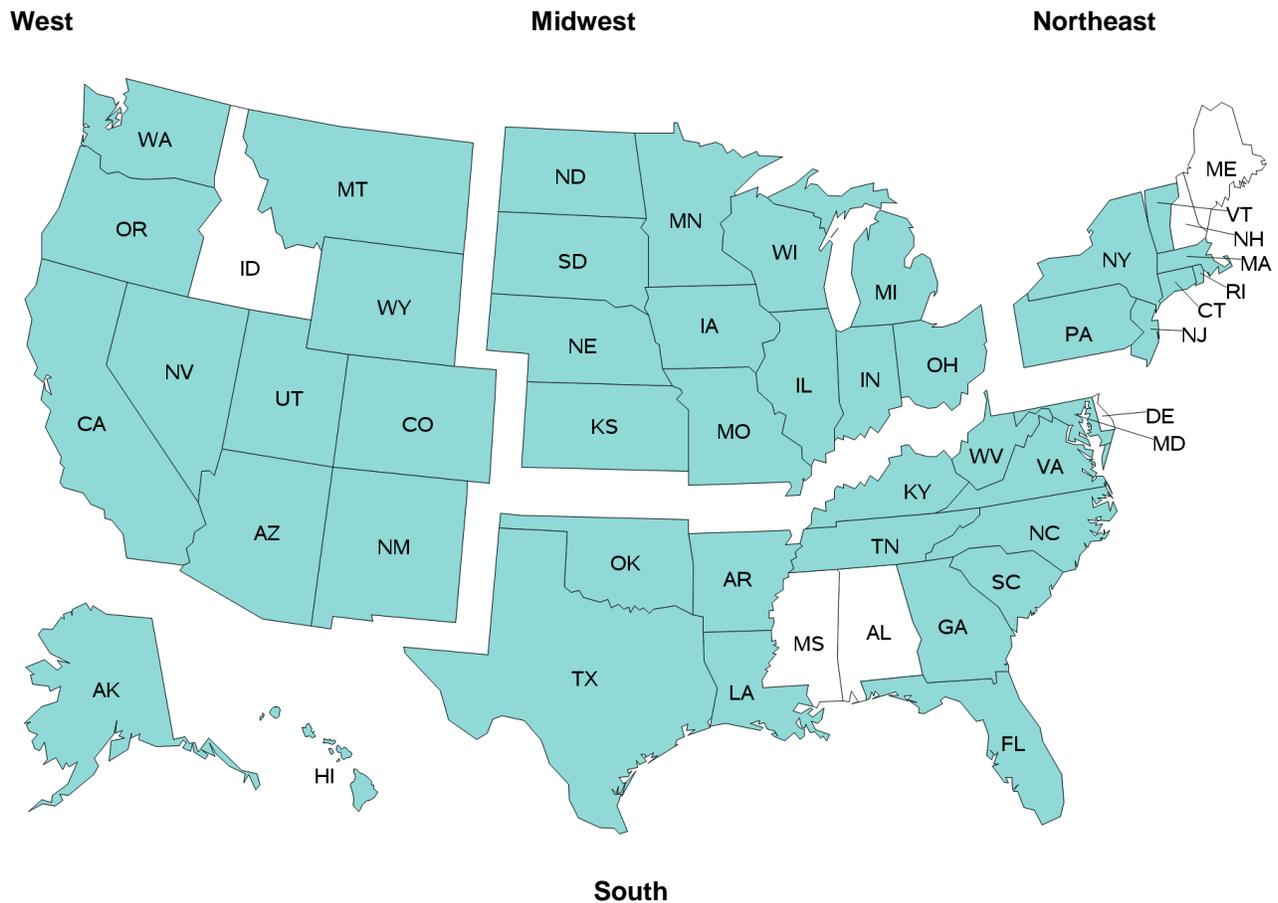
Table 2. Summary of KID States, Hospitals, and Inpatient Stays, 1997-2012

	2012	2009	2006	2003	2000	1997	
Number of States	44	44	38	36	27	22	
Data sources	AK AR AZ CA CO CT FL GA HI IA IL IN KS KY LA MA MD MI MN MO MT NC ND NE NJ NM NV NY OH OK OR PA RI SC SD TN TX UT VA VT WA WI WV WY (Added AK, ND. ME and NH are not included)	AR AZ CA CO CT FL GA HI IA IL IN KS KY LA MA ME MD MI MN MO MT NC NE NH NM NJ NV NY OH OK OR PA RI SC SD TN TX UT VA VT WA WI WV WY (Added LA, ME, MT, NM, PA and WY)	AR AZ CA CO CT FL GA HI IA IL IN KS KY MA MD MI MN MO NC NE NH NJ NV NY OH OK OR RI SC SD TN TX UT VA VT WA WI WV (Added AR and OK. ME and PA are not included)	AZ CA CO CT FL GA HI IA IL IN KS KY MD MA MI MN MO NC NE NH NJ NV NY OH OR RI SC SD TN TX UT VA VT WA WI WV (Added IL, IN, MI, MN, NE, NH, NV, OH, RI, SD, VT. ME and PA are not included)	AZ CA CO CT FL GA HI IA KS KY MD MA ME MO NC NJ NY OR PA SC TN TX UT VA WA WI WV (Added KY, ME, NC, TX, VA, WV. IL is not included)	AZ CA CO CT FL GA HI IL IA KS MD MA MO NJ NY OR PA SC TN UT WA WI	Community hospitals, including rehabilitation hospitals
Hospitals	Community, non-rehabilitation hospitals					Community hospitals, including rehabilitation hospitals	
Hospital universe¹⁰	5,118	5,128	5,124	4,836	4,839	5,113	
Number of hospitals with pediatric discharges	4,179	4,121	3,739	3,438	2,784	2,521	
Definition of pediatric discharges	Age at admission of 20 years or less for all years after 1997					Age at admission of 18 years or less	
Number of pediatric discharges (unweighted)	3,195,782	3,407,146	3,131,324	2,984,129	2,516,833	1,905,797	
Number of pediatric discharges (weighted)	6,675,222	7,370,203	7,558,812	7,409,162	7,291,032	6,657,326	

[Return to Introduction](#)

¹⁰ Most AHA survey responses from hospitals cover a fiscal year other than a January-to-December calendar year. The numbers of hospitals for the KID are based on the AHA Annual Survey files.

Figure 1. KID States, by Region, 2012¹¹



All States, by Region¹²

Region	States
1: Northeast	Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, Vermont.
2: Midwest	Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, Wisconsin.
3: South	<i>Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, West Virginia.</i>
4: West	Alaska, Arizona, California, Colorado, Hawaii, <i>Idaho</i> , Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming.

[Return to Introduction](#)

¹¹ New Hampshire, Maine, and Mississippi participate in HCUP, but did not provide data in time for the 2012 KID.

¹² States and areas in italics do not participate in HCUP.

Table 3. Summary of KID Releases

Data from	Media/format options	Structure of Releases
<ul style="list-style-type: none"> ▪ 1997 ▪ 22 States 	On CD-ROM, in ASCII format	1 year of data on one CD, compressed files
<ul style="list-style-type: none"> ▪ 2000 ▪ 27 States 		Beginning in 2003, a companion file with four different sets of severity measures
<ul style="list-style-type: none"> ▪ 2003 ▪ 36 States 	On DVD-ROM, In ASCII format	Beginning in 2006, a companion file with diagnosis and procedure groups
<ul style="list-style-type: none"> ▪ 2006 ▪ 38 States 		Beginning in 2009, 1 year of data in ASCII format on a single DVD-ROM
<ul style="list-style-type: none"> ▪ 2009 ▪ 44 States 	On DVD-ROM, In ASCII format	
<ul style="list-style-type: none"> ▪ 2012 ▪ 44 States 		

[Return to Introduction](#)

Table 4. KID Related Reports and Database Documentation Available on HCUP-US

<p>Restrictions on the Use of the KID</p> <ul style="list-style-type: none"> • Data Use Agreement for the KID 	<p>Known Data Issues</p> <ul style="list-style-type: none"> • Information on corrections to the KID data sets • Link to KID Trends Weights Files
<p>Description of the KID Files</p> <ul style="list-style-type: none"> • Introduction to the KID, 2012 – <i>this document</i> • HCUP Quality Control Procedures – describes procedures used to assess data quality • File Specifications – details data file names, number of records, record length, and record layout • Sources of KID Data and State-Specific Restrictions (<i>included in this document beginning 2006</i>) – identifies the KID data sources and restrictions on sampling and the release of data elements 	<p>Load Programs Programs to load the ASCII data files into statistical software:</p> <ul style="list-style-type: none"> • SAS • SPSS • Stata
<p>Availability of Data Elements</p> <ul style="list-style-type: none"> • Availability of KID data elements from 1997-2012 	<p>HCUP Tools: Labels and Formats</p> <ul style="list-style-type: none"> • Overview of Clinical Classifications Software (CCS), a categorization scheme that groups ICD-9-CM diagnosis and procedure codes into mutually exclusive categories • Label file for CCS categories • Label file for multiple versions of Diagnosis Related Groups (DRGs) and Major Diagnostic Categories (MDC) • KID SAS format library program to create value labels
<p>Description of Data Elements in the KID</p> <ul style="list-style-type: none"> • Description of Data Elements – details uniform coding and State-specific idiosyncrasies • Summary Statistics – lists means and frequencies on nearly all data elements • KID Severity Measures – provides detailed documentation on the different types of measures • HCUP Coding Practices – describes how HCUP data elements are coded • HCUP Hospital Identifiers – explains data elements that characterize individual hospitals 	<p>KID Related Reports Links to HCUP-US page with various KID related reports such as the following:</p> <ul style="list-style-type: none"> • Design of the Kids' Inpatient Databases for 1997, 2000 and 2003 (<i>included in this document beginning 2006</i>) • Changes in NIS Sampling and Weighting Strategy for 1998 (<i>for information on the 1998 NIS universe and strata which the KID uses</i>) • Calculating KID Variances • File Composition by State • KID Trends Report • KID Comparison Reports • HCUP E-Code Evaluation Report <p>HCUP Supplemental Files</p> <ul style="list-style-type: none"> • Cost-to-Charge Ratio files • Hospital Market Structure files • KID Trends Supplemental File

[Return to Introduction](#)

Table 5. Bed Size Categories, by Region

Location and Teaching Status	Hospital Bed Size		
	Small	Medium	Large
NORTHEAST			
Rural	1-49	50-99	100+
Urban, non-teaching	1-124	125-199	200+
Urban, teaching	1-249	250-424	425+
MIDWEST			
Rural	1-29	30-49	50+
Urban, non-teaching	1-74	75-174	175+
Urban, teaching	1-249	250-374	375+
SOUTH			
Rural	1-39	40-74	75+
Urban, non-teaching	1-99	100-199	200+
Urban, teaching	1-249	250-449	450+
WEST			
Rural	1-24	25-44	45+
Urban, non-teaching	1-99	100-174	175+
Urban, teaching	1-199	200-324	325+

[Return to Introduction](#)

Table 6: Number of AHA, HCUP SID, and KID Hospitals, by Region (the KID Stratifier) and by Census Division, 2012

Census Region / Division	AHA Universe Hospitals*	SID Community Non-Rehab Hospitals	SID Community Non-Rehab Hospitals With Peds Discharges	KID Sampling Frame Hospitals	KID Sample Hospitals
Total	5,118	4,540	4,213	4,213	4,179
Census Region					
1: Northeast	629	544	522	522	518
2: Midwest	1,500	1,381	1,322	1,322	1,310
3: South	2,031	1,742	1,534	1,534	1,520
4: West	958	873	835	835	831
Census Division					
1: New England	197	118	117	117	115
2: Middle Atlantic	432	426	405	405	403
3: East North Central	790	737	714	714	710
4: West North Central	710	644	608	608	600
5: South Atlantic	741	715	686	686	683
6: East South Central	432	311	204	204	202
7: West South Central	858	716	644	644	635
8: Mountain	408	339	321	321	321
9: Pacific	550	534	514	514	510

*The columns in the table are defined as follows:

- “AHA Universe Hospitals” lists all community, non-rehabilitation hospitals in the AHA Survey data.
- “SID Community, Non-Rehabilitation Hospitals” lists potential KID sampling-frame hospitals before applying restrictions to the frame and before excluding hospitals without any pediatric discharges.
- “SID Community, Non-Rehabilitation Hospitals with Pediatric Discharges” lists potential KID sampling-frame hospitals with pediatric discharges before applying restrictions to the frame.
- “KID Sampling-Frame Hospitals” lists hospitals with pediatric discharges in the sampling frame after applying state-specific restrictions to the frame.
- “KID Sample Hospitals” lists the hospitals selected for the KID. Some hospitals may not be included in the KID because they had so few pediatric discharges that none were randomly sampled.

[Return to Introduction](#)

Table 7. Number of Hospitals and Discharges in the AHA Universe, SID, and KID, by Hospital Type, 2012

	AHA Universe		SID		KID	
Hospital Type	Hospitals	Admissions Plus Births	Hospitals with Pediatric Discharges	Pediatric Discharges	Hospitals	Pediatric Discharges
Children's Hospital	80	621,259	70	505,999	70	400,835
Not a Children's Hospital	5,038	37,432,053	4,143	5,681,148	4,109	2,794,947
Total	5,118	38,053,312	4,213	6,187,147	4,179	3,195,782

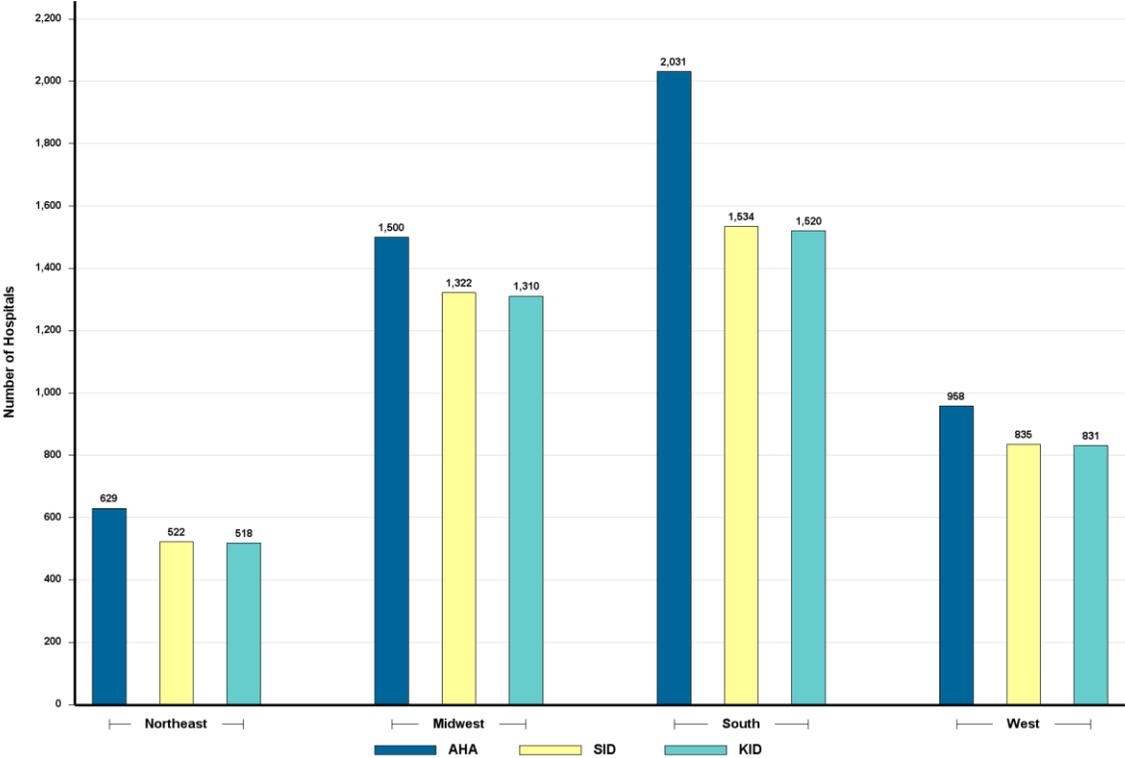
[Return to Introduction](#)

Table 8. 2012 KID Discharges, by Hospital Type

Hospital Type	Uncomplicated Births	Complicated Births	Pediatric Non-Births	Total Pediatric Discharges
Unweighted:				
Children's Hospital	566	3,921	396,348	400,835
Not a Children's Hospital	249,994	853,092	1,691,861	2,794,947
Total	250,560	857,013	2,088,209	3,195,782
Weighted:				
Children's Hospital	6,996	6,055	586,499	599,550
Not a Children's Hospital	2,610,848	1,109,861	2,354,963	6,075,672
Total	2,617,844	1,115,916	2,941,462	6,675,222

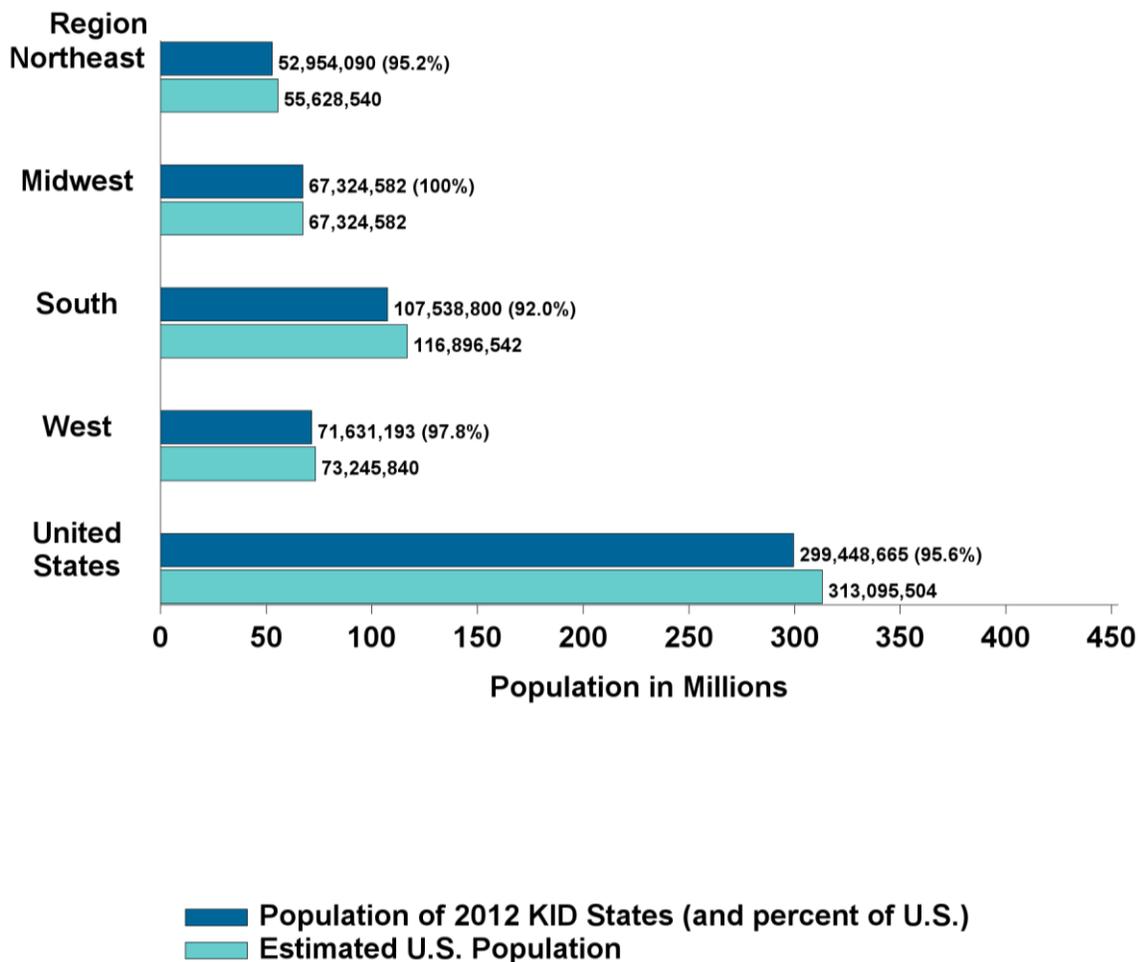
[Return to Introduction](#)

Figure 2. Number of Hospitals in the 2012 AHA Universe, SID, and KID, by Region



[Return to Introduction](#)

Figure 3. Percentage of U.S. Population in 2012 KID States, by Region
 Calculated using the estimated U.S. population on July 1, 2012.¹³



[Return to Introduction](#)

¹³ Table 1. Annual Estimates of the Population for the United States, Regions, States, and Puerto Rico: April 1, 2010 to July 1, 2013 (NST-EST2013-01). Source: U.S. Census Bureau, Population Division. Release Date: December 2013.

Appendix II: State-Specific Restrictions

The table below enumerates the types of restrictions applied to the KID. Restrictions include the following types:

- [Confidentiality of records](#)
 - Restricted release of age in years
- [Missing discharges](#).

Confidentiality of Records - Restricted Release of Age in Years

- At least one Partner required ages in years (AGE) to be set to the midpoints of age ranges.

Missing Discharges

- At least one Partner prohibits the release of discharge records for patients with HIV diagnoses.
- At least one Partner prohibits the release of behavioral health including chemical dependency care or psychiatric care discharges.
- At least one Partner prohibits the release of Alternate Level of Care (SNF / Swing Bed Skilled) discharges.
- At least one Partner prohibits the release of abortion discharges.

[Return to Introduction](#)

Appendix III: Data Elements

Table 1. Data Elements in the 2012 KID Inpatient Core File

For prior years, refer to documentation on HCUP-US (e.g. the table of data element availability by years http://hcup-us.ahrq.gov/db/nation/kid/Availability_of_KID_Data_Elements_2012.pdf or previous versions of the KID Introduction).

Type of Data Element	HCUP Name	Coding Notes
Admission day of week or weekend	AWEEKEND	Admission on weekend: (0) admission on Monday-Friday, (1) admission on Saturday-Sunday
Admission month	AMONTH	Admission month coded from (1) January to (12) December
Transferred into hospital	TRAN_IN	Transfer In Indicator: (0) not a transfer, (1) transferred in from a different acute care hospital [ATYPE NE 4 & (ASOURCE=2 or POO=4)], (2) transferred in from another type of health facility [ATYPE NE 4 & (ASOURCE=3 or POO=5, 6)]
Admission type	ELECTIVE	Indicates elective admission: (1) elective, (0) non-elective admission
Age at admission	AGE	Age in years coded 0-124 years
Chronic Conditions	NCHRONIC	Number of chronic conditions
Clinical Classifications Software (CCS) category	DXCCS1 – DXCCS25	CCS category for all diagnoses. Beginning in 2009, the diagnosis array was increased from 15 to 25.
	PRCCS1 – PRCCS15	CCS category for all procedures
Diagnosis information	DX1 – DX25	Diagnoses, principal and secondary (ICD-9-CM). Beginning in 2003, the diagnosis array does not include any of external cause of injury codes. These codes have been stored in a separate array ECODEn. Beginning in 2009, the diagnosis array was increased from 15 to 25.
	HOSPBIRTH	Birth diagnosis, in this hospital
	NDX	Number of diagnoses coded on the original record
	UNCBIRTH	Normal, uncomplicated birth in hospital
Diagnosis Related Group (DRG)	DRG	DRG in use on discharge date
	DRG_NoPOA	DRG in use on discharge date, calculated without Present On Admission (POA) indicators
	DRGVER	Group version in use on discharge date
	DRG24	DRG Version 24 (effective October 2006 - September 2007)
Discharge quarter	DQTR	Coded: (1) Jan - Mar, (2) Apr - Jun, (3) Jul - Sep, (4) Oct - Dec

Type of Data Element	HCUP Name	Coding Notes
Discharge weights	DISCWT	Weight to discharges in AHA universe for national estimates. In 2000, the discharge weight DISCWTCHARGE should be used for estimates of total charges.
Discharge year	YEAR	Calendar year
Disposition of patient (discharge status)	DIED	Indicates in-hospital death: (0) did not die during hospitalization, (1) died during hospitalization
	DISPUNIFORM	Disposition of patient, uniform coding used beginning in 1998: (1) routine, (2) transfer to short term hospital, (5) other transfers, including skilled nursing facility, intermediate care, and another type of facility, (6) home health care, (7) against medical advice, (20) died in hospital, (99) discharged alive, destination unknown
	TRAN_OUT	Transfer Out Indicator: (0) not a transfer, (1) transferred out to a different acute care hospital, (2) transferred out to another type of health facility
External causes of injury and poisoning	ECODE1 – ECODE4	External cause of injury and poisoning code, primary and secondary (ICD-9-CM). Beginning in 2003, external cause of injury codes are stored in a separate array ECODEn from the diagnosis codes in the array DXn. Prior to 2003, these codes are contained in the diagnosis array (DXn).
	E_CCS1 – E_CCS4	CCS category for the external cause of injury and poisoning codes
	NECODE	Number of external cause of injury codes on the original record.
Gender of patient	FEMALE	Indicates gender for KID beginning in 1998: (0) male, (1) female
Hospital information	HOSP_REGION	Region of hospital: (1) Northeast, (2) Midwest, (3) South, (4) West Prior to 2012, region of hospital is only available in the KID Hospital File.
	KID_STRATUM	Hospital stratum used for weights.
Indicates Emergency Department service	HCUP_ED	Indicator that discharge record includes evidence of emergency department (ED) services: (0) Record does not meet any HCUP Emergency Department criteria, (1) Emergency Department revenue code on record, (2) Positive Emergency Department charge (when revenue center codes are not available), (3) Emergency Department CPT procedure code on record, (4) Admission source of ED, (5) State-defined ED record; no ED charges available
Length of Stay	LOS	Length of stay, edited
Location of the patient	PL_NCHS2006	Urban–rural designation for patient's county of residence: (1) "Central" counties of metro areas >= 1 million population, (2) "Fringe" counties of metro areas >= 1 million population, (3) Counties in metro areas of 250,000 - 999,999 population, (4) Counties in metro areas of 50,000 - 249,999 population, (5) micropolitan counties, (6) non-core counties

Type of Data Element	HCUP Name	Coding Notes
Major Diagnosis Category (MDC)	MDC	MDC in use on discharge date
	MDC_NoPOA	MDC in use on discharge date, calculated without Present on Admission (POA) indicators
	MDC24	MDC Version 24 (effective October 2006 - September 2007)
Median household income for patient's ZIP Code	ZIPINC_QRTL	Median household income quartiles for patient's ZIP Code. Because these estimates are updated annually, the value ranges for the ZIPINC_QRTL categories vary by year. Check the HCUP-US Website for details.
Neonatal/maternal flag	NEOMAT	Assigned from diagnoses and procedure codes: (0) not maternal or neonatal, (1) maternal diagnosis or procedure, (2) neonatal diagnosis, (3) maternal and neonatal on same record
Payer information	PAY1	Expected primary payer, uniform: (1) Medicare, (2) Medicaid, (3) private including HMO, (4) self-pay, (5) no charge, (6) other
Procedure information	PR1 – PR15	Procedures, principal and secondary (ICD-9-CM)
	NPR	Number of procedures coded on the original record
	ORPROC	Major operating room procedure indicator: (0) no major operating room procedure, (1) major operating room procedure
	PRDAY1	Number of days from admission to principal procedure.
	PRDAY2 – PRDAY15	Number of days from admission to secondary procedures.
Race of Patient	RACE ¹⁴	Race, uniform coding: (1) white, (2) black, (3) Hispanic, (4) Asian or Pacific Islander, (5) Native American, (6) other
Record identifier, synthetic	RECNUM	HCUP unique record number
Total Charges	TOTCHG	Total charges, edited

[Return to Introduction](#)

¹⁴ Race contains missing values on more than 8% of the records.

Table 2. Data Elements in the 2012 KID Hospital File

For prior years, refer to documentation on HCUP-US (e.g. the table of data element availability by years http://hcup-us.ahrq.gov/db/nation/kid/Availability_of_KID_Data_Elements_2012.pdf or previous versions of the KID Introduction).

Type of Data Element	HCUP Name	Coding Notes
Universe Counts	N_DISC_U	Number of universe discharges in the KID_STRATUM
	N_BRTH_U	Number of universe births in KID_STRATUM
	N_HOSP_U	Number of universe hospitals in KID_STRATUM
Sample Counts	S_DISC_U	Number of sampled discharges in the sampling stratum (KID_STRATUM or STRATUM)
	S_BRTH_U	Number of sample births in KID_STRATUM
	S_CHLD_U	Number of sample pediatric non-births in KID_STRATUM
	S_CMPB_U	Number of sample complicated births in KID_STRATUM
	S_UNCB_U	Number of sample uncomplicated births in KID_STRATUM
	S_HOSP_U	Number of sample hospitals in KID_STRATUM
Hospital Characteristics	KID_STRATUM	Hospital stratum used for weights
	HOSP_BEDSIZE	Bed size of hospital (STRATA): (1) small, (2) medium, (3) large
	H_CONTRL	Control/ownership of hospital (STRATA): (1) government, nonfederal, (2) private, non-profit, (3) private, invest-own
	HOSP_LOCTEACH	Location/teaching status of hospital (STRATA): (1) rural, (2) urban non-teaching, (3) urban teaching
	HOSP_REGION	Region of hospital (STRATA): (1) Northeast, (2) Midwest, (3) South, (4) West
	HOSP_DIVISION	Census Division of hospital: (1) New England, (2) Middle Atlantic, (3) East North Central, (4) West North Central, (5) South Atlantic, (6) East South Central, (7) West South Central, (8) Mountain, (9) Pacific
Discharge Year	YEAR	Calendar year

[Return to Introduction](#)

Table 3. Data Elements in the 2012 KID Disease Severity Measures Files

For prior years, refer to documentation on HCUP-US (e.g. the table of data element availability by years http://hcup-us.ahrq.gov/db/nation/kid/Availability_of_KID_Data_Elements_2012.pdf or previous versions of the KID Introduction).

Type of Data Element	HCUP Name	Coding Notes
AHRQ Comorbidity Software (AHRQ)	CM_AIDS	AHRQ comorbidity measure: Acquired immune deficiency syndrome
	CM_ALCOHOL	AHRQ comorbidity measure: Alcohol abuse
	CM_ANEMDEF	AHRQ comorbidity measure: Deficiency anemias
	CM_ARTH	AHRQ comorbidity measure: Rheumatoid arthritis/collagen vascular diseases
	CM_BLDLOSS	AHRQ comorbidity measure: Chronic blood loss anemia
	CM_CHF	AHRQ comorbidity measure: Congestive heart failure
	CM_CHRNLUNG	AHRQ comorbidity measure: Chronic pulmonary disease
	CM_COAG	AHRQ comorbidity measure: Coagulopathy
	CM_DEPRESS	AHRQ comorbidity measure: Depression
	CM_DM	AHRQ comorbidity measure: Diabetes, uncomplicated
	CM_DMCX	AHRQ comorbidity measure: Diabetes with chronic complications
	CM_DRUG	AHRQ comorbidity measure: Drug abuse
	CM_HTN_C	AHRQ comorbidity measure: Hypertension, uncomplicated and complicated
	CM_HYPOTHY	AHRQ comorbidity measure: Hypothyroidism
	CM_LIVER	AHRQ comorbidity measure: Liver disease
	CM_LYMPH	AHRQ comorbidity measure: Lymphoma
	CM_LYTES	AHRQ comorbidity measure: Fluid and electrolyte disorders
	CM_METS	AHRQ comorbidity measure: Metastatic cancer
	CM_NEURO	AHRQ comorbidity measure: Other neurological disorders
	CM_OBESE	AHRQ comorbidity measure: Obesity
	CM_PARA	AHRQ comorbidity measure: Paralysis
	CM_PERIVASC	AHRQ comorbidity measure: Peripheral vascular disorders
	CM_PSYCH	AHRQ comorbidity measure: Psychoses
CM_PULMCIRC	AHRQ comorbidity measure: Pulmonary circulation disorders	
CM_RENLFAIL	AHRQ comorbidity measure: Renal failure	
CM_TUMOR	AHRQ comorbidity measure: Solid tumor without metastasis	
CM_ULCER	AHRQ comorbidity measure: Peptic ulcer disease excluding bleeding	

Type of Data Element	HCUP Name	Coding Notes
	CM_VALVE	AHRQ comorbidity measure: Valvular disease
	CM_WGHTLOSS	AHRQ comorbidity measure: Weight loss
All Patient Refined DRG (3M)	APRDRG	All Patient Refined DRG
	APRDRG_Risk_Mortality	All Patient Refined DRG: Risk of Mortality Subclass
	APRDRG_Severity	All Patient Refined DRG: Severity of Illness Subclass
Linkage Variables	HOSP_KID	KID hospital number (links to Hospital Weights file; does not link to previous years)
	RECNUM	HCUP record identifier (links to KID discharge level files; does not link to previous years)

[Return to Introduction](#)

Table 4. Data Elements in the 2012 KID Diagnosis and Procedure Groups Files

For prior years, refer to documentation on HCUP-US (e.g. the table of data element availability by years http://hcup-us.ahrq.gov/db/nation/kid/Availability_of_KID_Data_Elements_2012.pdf or previous versions of the KID Introduction).

Type of Data Element	HCUP Name	Coding Notes
Chronic Condition Indicator	CHRON1 – CHRON25	Chronic condition indicator for all diagnoses: (0) non-chronic condition, (1) chronic condition. Beginning in 2009, the diagnosis array was increased from 15 to 25.
	CHRONB1 – CHRONB25	Body system for all diagnoses: (1) Infectious and parasitic disease, (2) Neoplasms, (3) Endocrine, nutritional, and metabolic diseases and immunity disorders, (4) Diseases of blood and blood-forming organs, (5) Mental disorders, (6) Diseases of the nervous system and sense organs, (7) Diseases of the circulatory system, (8) Diseases of the respiratory system, (9) Diseases of the digestive system, (10) Diseases of the genitourinary system, (11) Complications of pregnancy, childbirth, and the puerperium, (12) Diseases of the skin and subcutaneous tissue, (13) Diseases of the musculoskeletal system, (14) Congenital anomalies, (15) Certain conditions originating in the perinatal period, (16) Symptoms, signs, and ill-defined conditions, (17) Injury and poisoning, (18) Factors influencing health status and contact with health services. Beginning in 2009, the diagnosis array was increased from 15 to 25.
Multi-Level CCS: Principal Diagnosis	DXMCCS1	Multi-level clinical classification software (CCS) for principal diagnosis. Four levels for diagnoses presenting both the general groupings and very specific conditions
Multi-Level CCS: E Code 1	E_MCCS1	Multi-level clinical classification software (CCS) for first listed E Code. Four levels for E codes presenting both the general groupings and very specific conditions
Multi-Level CCS: Principal Procedure	PRMCCS1	Multi-level clinical classification software (CCS) for principal procedure. Three levels for procedures presenting both the general groupings and very specific conditions
Procedure Class	PCLASS1 – PCLASS15	Procedure Class for all procedures: (1) Minor Diagnostic, (2) Minor Therapeutic, (3) Major Diagnostic, (4) Major Therapeutic
Linkage Variables	HOSP_KID	KID hospital number (links to Hospital Weights file; does not link to previous years)
	RECNUM	HCUP record identifier

[Return to Introduction](#)

Appendix IV: Teaching Hospital Indicator Assignment

We used the following American Hospital Association Annual Survey Database (Health Forum, LLC © 2013) data elements to assign the KID Teaching Hospital Indicator:

AHA Data Element Name = Description [HCUP Data Element Name].

BDH = Number of short-term hospital beds [B001H].
BDTOT = Number of total facility beds [B001].
FTRES = Number of full-time employees: interns & residents (medical & dental) [E125].
PTRES = Number of part-time employees: interns & residents (medical & dental) [E225].
MAPP8 = Council of Teaching Hospitals (COTH) indicator [A101].
MAPP3 = Residency training approval by the Accreditation Council for Graduate Medical Education (ACGME) [A102].

Prior to the 1998 KID, we used the following SAS code to assign the KID teaching hospital status indicator, H_TCH:

```
/* FIRST ESTABLISH SHORT-TERM BEDS DEFINITION */
IF BDH NE . THEN BEDTEMP = BDH ; /* SHORT TERM BEDS */
ELSE IF BDH =. THEN BEDTEMP=BDTOT ; /* TOTAL BEDS PROXY */

/*****/
/* NEXT ESTABLISH TEACHING STATUS BASED ON F-T & P-T */
/* RESIDENT/INTERN STATUS FOR HOSPITALS. */
/*****/
RESINT = (FTRES + .5*PTRES)/BEDTEMP ;
IF RESINT > 0 & (MAPP3=1 OR MAPP8=1) THEN H_TCH=1; /* 1=TEACHING */
ELSE H_TCH=0 ; /* 0=NONTEACHING */
```

Beginning with the 1998 KID, we used the following SAS code to assign the teaching hospital status indicator, HOSP_TEACH:

```
/******  
/* FIRST ESTABLISH SHORT-TERM BEDS DEFINITION */  
/******  
IF BDH NE . THEN BEDTEMP = BDH ; /* SHORT TERM BEDS */  
ELSE IF BDH =. THEN BEDTEMP = BDTOT ; /* TOTAL BEDS PROXY */  
/******  
/* ESTABLISH IRB NEEDED FOR TEACHING STATUS */  
/* BASED ON F-T P-T RESIDENT INTERN STATUS */  
/******  
IRB = (FTRES + .5*PTRES) / BEDTEMP ;  
/******  
/* CREATE TEACHING STATUS DATA ELEMENT */  
/******  
IF (MAPP8 EQ 1) OR (MAPP3 EQ 1) THEN HOSP_TEACH = 1 ;  
ELSE IF (IRB GE 0.25) THEN HOSP_TEACH = 1 ;  
ELSE HOSP_TEACH = 0 ;
```

[Return to Introduction](#)