

An Introduction to the Commonwealth Institute of Kentucky's HCUP-KID Data

Md Yasin Ali Parh¹, MS; Seyed Karimi¹, PhD

¹University of Louisville, School of Public Health and Information Sciences

1. Background

The Agency for Healthcare Research and Quality's (AHRQ) Healthcare Cost and Utilization Project (HCUP) data is the biggest longitudinal hospital care data collection in the United States (U.S.). HCUP is essentially a family of healthcare databases that started in 1988. The databases are:

- National Inpatient Sample (NIS),
- Kids' Inpatient Database (KID),
- Nationwide Ambulatory Surgery Sample (NASS),
- Nationwide Emergency Department Sample (NEDS),
- Nationwide Readmissions Database (NRD),
- State Inpatient Databases (SID),
- State Ambulatory Surgery and Services Databases (SASD), and
- State Emergency Department Databases (SEDD)

See <https://www.hcup-us.ahrq.gov> for more information. These databases are built based on information from state data organizations, hospital associations, private data organizations, and the federal government. Together with HCUP software tools and products, they can be used for research on a variety of contemporary healthcare topics and trends. Below is a brief overview of the National Readmission Database (NRD) data.

The KID is a part of the HCUP that is available every 3 years starting from 1997. Hospital inpatient stays of children were used in the creation of this largest publicly available pediatric inpatient care database. This database is a representative sample of pediatric discharges from community (non-rehabilitation) hospitals in 36 states. Patients' age in the data is 20 years or younger, and the charge information from all expected payers is included. The KID records around 2-3 million pediatric inpatient stays annually from 2,500-4,200 hospitals. Inpatient stay records and hospital data with weights are included in every KID release in order to create national estimates (See: <https://hcup-us.ahrq.gov/kidoverview.jsp>).

2. Database information, documentation, and purchasing information:

The KID database is designed to investigate a broad spectrum of topics, including the financial cost of pediatric illnesses, access to services, the quality of care and patient safety, and the impact of health policy changes. Details about this public use databases and documentation are available at this <https://hcup-us.ahrq.gov/db/nation/kid/kidarchive.jsp>.

All purchasers and users of KID data must complete the web-based HCUP data use agreement (DUA) training course and sign an HCUP DUA before receipt of the data. HCUP data use agreement, training, and responsibility of HCUP data purchaser are available at this https://hcup-us.ahrq.gov/tech_assist/dua.jsp.

3. KID Data Elements and Summary Statistics

Data element descriptions explain how the data element is coded in the HCUP databases and what are the uniform values. The descriptions are cumulative across all years of KID data from 1997 to the current data year. Detailed data element information is available at <https://hcup-us.ahrq.gov/db/nation/kid/kiddde.jsp>, and each year record layouts for the KID file are available at this <https://hcup-us.ahrq.gov/db/nation/kid/kidfilespecs.jsp>. The KID can be used for analyses of rare conditions, such as congenital anomalies and uncommon treatments, such as cardiac surgery. For each KID data, a summary statistic is available at <https://hcup-us.ahrq.gov/db/nation/kid/kidsummarystats.jsp>. Summary statistics include means on all numeric variables, frequency distribution, and univariates on continuous variables.

The recent (2019) KID data recorded 3,089,283 inpatient stays with a weighted sum of 5,902,538 records and 15,606 patients who died during hospitalization. The average age of the patients was 5.6 years, where most (55.06%) of the kids' age were less than 1 year. The proportion of male and female kids was the same. In 2019, the average length of hospital stay was 5.0 days. Most of the patients were white (45.57%), 19.66% Hispanic, 16.63% black, 6.09% other races, 4.0% Asian/Pacific Islander, 0.85% Native American, and the rest of the patient's race were missing.

4. Commonwealth Institute of Kentucky's (CIK) NIS Data

4.1. Years of Data

The CIK and the UofL Department of Pediatrics Child and Adolescent Health Research Design and Support Unit (CAHRDS) have jointly purchased the triannual KID datasets from 1997 to 2019.

The data, which are in ASCII format, include

- KID core file
- KID hospital file

- KID diagnosis and procedure groups file, and
- KID severity measures file

The following describes the contents/variables of these four data files. The exact name of each variable in the data is also provided and hyperlinked to the HUCP webpage for documentation and summary statistics information.

4.2. Components of the 2018 NIS Core File

- **Patient demographic and location information**
 - Age in years at admission ([AGE](#))
 - Neonatal age (first 28 days after birth) indicator ([AGE_NEONATE](#))
 - Indicator of sex ([FEMALE](#))
 - Race ([RACE](#))
 - Patient Location: NCHS urban-rural code ([PL_NCHS](#))
 - Median household income national quartile for the patient ([ZIPINC_QRTL](#))
- **Admission information**
 - If admission day was a weekend ([AWEEKEND](#))
 - Admission month ([AMONTH](#))
 - Transfer in indicator ([TRAN_IN](#))
 - HCUP emergency department service indicator ([HCUP_ED](#))
 - Elective versus non-elective admission ([ELECTIVE](#))
- **Diagnosis and procedure information**
 - ICD-10-CM diagnosis ([I10_DXn](#)); up to 40 diagnoses; n = 1, 2, ..., 40
 - Number of ICD-10-CM diagnoses on this record ([I10_NDX](#))
 - ICD-10-PCS Procedure ([I10_PRn](#)); up to 25 procedures; n = 1, 2, ..., 25
 - ICD-10-CM diagnosis indicates in-hospital birth ([I10_HOSPBIRTH](#))
 - ICD-10-CM diagnosis indicates a normal uncomplicated in-hospital birth ([I10_UNCBIRTH](#))
 - Number of ICD-10-PCS procedures on this record ([I10_NPR](#))
 - Number of days from admission to principal procedure ([PRDAY1](#))
 - Number of days from admission to secondary procedures ([PRDAYn](#)); up to 24 procedures on admission; n = 2, ..., 25
- **Major diagnostic categories (MDCs)**
 - MDC in effect on discharge date ([MDC](#))
 - MDC in use on discharge date, calculated without present on admission, POA ([MDC_NoPOA](#))
- **Diagnosis-related group (DRG)**
 - Diagnosis-related group (DRG) in effect on discharge date ([DRG](#))
 - DRG grouper version used on discharge date ([DRGVER](#))

- DRG in use on discharge date, calculated without POA ([DRG_NoPOA](#))
- **Indicators**
 - ICD-10-CM Birth Indicator ([I10_BIRTH](#))
 - ICD-10-CM Delivery Indicator ([I10_DELIVERY](#))
 - Injury ICD-10-CM diagnosis reported on record ([I10_INJURY](#))
 - Multiple ICD-10-CM injuries reported on record ([I10_MULTINJURY](#))
 - Service line based on ICD-10-CM/PCS codes ([I10_SERVICELINE](#))
 - Major operating room ICD-10-PCS procedure indicator ([PCLASS_ORPROC](#))
- **Resource use information**
 - Total charges ([TOTCHG](#))
 - Length of stay ([LOS](#))
- **Discharge information**
 - Discharge quarter ([DQTR](#))
 - Calendar year ([YEAR](#))
 - Died during hospitalization ([DIED](#))
 - Disposition of patient ([DISPUNIFORM](#))
 - Transfer out indicator ([TRAN_OUT](#))
- **Linkage data element**
 - KID unique record number ([RECNUM](#))
 - * *RECNUM contains a unique HCUP record number that is specific to the Kids' Inpatient Database (KID). RECNUM can be used to link the KID Core and Severity Measures file*
 - KID hospital number ([HOSP_KID](#))
 - * *HOSP_KID cannot link to previous years or other databases.*
- **Weights (to calculate national estimates)**
 - KID discharge weight ([DISCWT](#))
- **Hospital information**
 - Census Division of the hospital ([HOSP_REGION](#))
 - KID hospital stratum ([KID_STRATUM](#))
- **Indicators**
 - ICD-10-CM Birth Indicator ([I10_BIRTH](#))
 - ICD-10-CM Delivery Indicator ([I10_DELIVERY](#))
 - Injury ICD-10-CM diagnosis reported on record ([I10_INJURY](#))
 - Multiple ICD-10-CM injuries reported on record ([I10_MULTINJURY](#))
 - Service line based on ICD-10-CM/PCS codes ([I10_SERVICELINE](#))
 - Major operating room ICD-10-PCS procedure indicator ([PCLASS_ORPROC](#))
- **Payer information**
 - Primary expected payer ([PAY1](#))

4.3. Components of the 2018 KID Hospital File

- **Universe counts**
 - Total number of inpatient discharges in the universe AHA community, excluding non-rehabilitation hospitals, in the stratum ([N_DISC_U](#))
 - Total number of hospitals in the universe of AHA community excluding non-rehabilitation hospital ([N_HOSP_U](#))
 - Number of births obtained from the AHA Annual Survey of Hospitals ([N_BRTH_U](#))
- **Sample counts**
 - Total number of inpatient discharges in sampled hospitals ([S_DISC_U](#))
 - Total Number of sample births in STRATUM ([S_BRTH_U](#))
 - Total number of sample pediatric non-births in STRATUM ([S_CHLD_U](#))
 - Number of complicated births sampled in the hospital ([S_CMPB_U](#))
 - Number of sample uncomplicated births in STRATUM ([S_UNCB_U](#))
 - Total number of sampled hospitals ([S_HOSP_U](#))
- **Hospital characteristics**
 - Bed size of the hospital ([HOSP_BEDSIZE](#))
 - Location/teaching status of the hospital ([HOSP_LOCTEACH](#))
 - Region of the hospital ([HOSP_REGION](#))
 - Control/ownership of hospital ([H_CONTRL](#))
 - KID hospital stratum ([KID_STRATUM](#))
- **Discharge information**
 - Calendar year ([Year](#))
- **Linkage data element**
 - KID hospital number ([HOSP_KID](#))
* *HOSP_KID cannot link to previous years or other databases.*

4.4. Components of the 2018 KID Severity File

- **Linkage data element**
 - KID unique record number ([RECNUM](#))
* *RECNUM contains a unique HCUP record number that is specific to the Kids' Inpatient Database (KID). RECNUM can be used to link the KID Core and Severity Measures file*
 - KID hospital number ([HOSP_KID](#))
* *HOSP_KID cannot link to previous years or other databases.*
- **All patients refined diagnosis-related group (DRG)**
 - All patient refined DRG ([APRDRG](#))
 - All patient refined DRG: risk of mortality subclass ([APRDRG_Risk_Mortality](#))
 - All patient refined DRG: severity of illness subclass ([APRDRG_Severity](#))

4.5. Components of the 2018 KID Diagnosis and Procedure Groups File

- **Linkage data element**
 - KID unique record number ([RECNUM](#))
 - KID hospital number ([HOSP_KID](#))
- **Clinical classifications software refined (CCSR) category**
 - At least one ICD-10-CM diagnosis on the record is included in CCSR *aaannn* ([DXCCSR_aaannn](#)¹)
 - Default clinical classifications software refined (CCSR) for principal diagnosis ([DXCCSR_DEFAULT_DX1](#))
 - Version of CCSR for ICD-10-CM diagnoses ([DXCCSR_VERSION](#))
 - Clinical Classifications Software Refined (CCSR): ICD-10-PCS Procedure Classification within Clinical Domain *aaa* ([PRCCSR_aaannn](#)²)
 - Version of the Clinical Classifications Software Refined (CCSR) for ICD-10-PCS procedure codes ([PRCCSR_VERSION](#))

¹*aaa* denotes the body system and *nnn* denotes the CCSR number within the body system.

²*aaa* denotes the clinical domain and *nnn* denotes the CCSR number within the clinical domain.

- **Procedure Classes Refined**
 - Procedure Classes Refined for ICD-10-PCS procedure codes ([PCLASSn](#)³)
 - Version of the Procedure Classes Refined for ICD-10-PCS procedure codes ([PCLASS_VERSION](#))

³PCLASSn was also available on the KID through data year 2012 and was specific to the coding of ICD-9-CM procedures.

Note: The data linkage element, KID hospital number, is available for all files. The KID record number is also available for all files except the hospital file.

5. Publications

The following link provides access to lists of publications on HCUP data.

<https://hcup-us.ahrq.gov/reports/pubsearch/pubsearch.jsp>