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

Md Yasin Ali Parh[1], MS; Seyed Karimi[1], PhD

[1]*University 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 Inpatient Sample (NIS) data.

The NIS has been a part of the HCUP from the beginning. NIS is the largest publicly available hospital inpatient stays database and is generated from the hospital billing information provided by statewide data organizations nationwide. In effect, NIS is sampled from the State Inpatient Databases (SID), including all inpatient data currently contributed to HCUP. Hospital stays information for all expected payers (i.e., self and private and public insurers) is reported in the data. Data from 49 statewide data organizations (48 States and the District of Columbia) are currently included in the NIS sampling frame, which covers approximately 98% of the U.S. population and roughly 97% of hospital discharges from community hospitals. Recent NIS records about 7 million inpatient stays annually (See: https://hcup-us.ahrq.gov/nisoverview.jsp).

COMMONWEALTH INSTITUTE OF KENTUCKY

2. Database information, documentation, and purchasing information:

The NIS database is designed to produce the U.S. regional and national estimates of inpatient utilization, access, cost, quality, and outcomes. Details about this public use databases and documentation are available at this https://hcup-us.ahrq.gov/db/nation/nis/nisarchive.jsp.

All purchasers and users of NIS 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 https://hcup-us.ahrq.gov/tech_assist/dua.jsp.

3. NIS Data Elements and Summary Statistics

Data element descriptions explain how the data elements are coded in the HCUP databases and what are the uniform values. The descriptions are cumulative across all years of NIS data from 1988 to the current data year. Detailed data element information is available at https://hcup-us.ahrq.gov/db/nation/nis/nisdde.jsp, and each year record layouts for the NIS file are available at https://hcup-us.ahrq.gov/db/nation/nis/nisfilespecs.jsp. For each year of NIS data, a summary statistic is available at this https://hcup-us.ahrq.gov/db/nation/nis/nissummstats.jsp. Summary statistics include means on all numeric variables, frequency distribution, and univariates on continuous variables. The recent (2020) NIS data recorded 6,471,165 inpatient stays with a weighted sum of 32,355,827 records. The inpatient stays include 180,360 patients who died during the hospitalization. The average age of the patients was 49.86 years (standard deviation, SD=27.18), and the majority of the hospital stayed patients were female (56%). Also, the average length of hospital stay was 4.94 days (SD=7.64). Most patients were white (62.46%), 14.66% black, 12.40% Hispanic, 3.47% other races, 3.02% Asian/Pacific Islander, 0.65% Native American, and the rest of the patients' 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 NIS datasets from 2000 to 2018.

The data, which are in ASCII format, include
- NIS core file
- NIS hospital file
- NIS diagnosis and procedure groups file, and
- NIS 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.

U **COMMONWEALTH**
INSTITUTE OF KENTUCKY

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 on 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
  - Number of ICD-10-PCS procedures on this record (I10_NPR)
  - Number of days from admission to procedure (PRDAYn); n = 1, 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)

- **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)

UoL **COMMONWEALTH INSTITUTE OF KENTUCKY**

- **Linkage data element**
  - NIS unique record number (KEY_NIS)
  - NIS hospital number (HOSP_NIS)

- **Weights (to calculate national estimates)**
  - NIS discharge weight (DISCWT)

- **Hospital information**
  - Census Division of the hospital (HOSP_DIVISION)
  - NIS hospital stratum (NIS_STRATUM)
- **Payer information**
  - Primary expected payer (PAY1)

4.3. **Components of the 2018 NIS Hospital File**

- **Linkage data element**
  - NIS hospital number (HOSP_NIS)

- **Weights (to calculate national estimates)**
  - NIS discharge weight (DISCWT)

- **Hospital characteristics**
  - Bed size of the hospital (HOSP_BEDSIZE)
  - Census Division of the hospital (HOSP_DIVISION)
  - Location/teaching status of the hospital (HOSP_LOCTEACH)
  - Region of the hospital (HOSP_REGION)
  - Control/ownership of hospital (H_CONTRL)
  - NIS hospital stratum (NIS_STRATUM)

- **Stratum information**
  - 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)
  - Total number of inpatient discharges in sampled hospitals (S_DISC_U)
  - Total number of sampled hospitals (S_HOSP_U)

- **Discharge information**
  - Total number of discharges from this hospital in the NIS (TOTAL_DISC)
  - Calendar year (Year)

Uof**L**® **COMMONWEALTH INSTITUTE OF KENTUCKY**

4.4. Components of the 2018 NIS Severity File

- **Linkage data element**
  - NIS unique record number (KEY_NIS)
  - NIS hospital number (HOSP_NIS)

- **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 NIS Diagnosis and Procedure Groups File

- **Linkage data element**
  - NIS unique record number (KEY_NIS)
  - NIS hospital number (HOSP_NIS)

- **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)

  [1]*aaa* denotes the body system, and *nnn* denotes the CCSR number within the body system. For example, the data element DXCCSR_CIR007 indicates the CCSR category for essential hypertension under the body system of diseases of the circulatory system.

Note: The data linkage element NIS hospital number is available for all files. The NIS record number is also available for all files except the hospital file. NIS hospital number, census division of the hospital, NIS hospital stratum, and calendar year are common in both core and hospital files.

5. Publications
The following link provides access to lists of publications on HCUP data.
https://hcup-us.ahrq.gov/reports/pubsearch/pubsearch.jsp

U of L COMMONWEALTH INSTITUTE OF KENTUCKY