

Programming with the HCUP SID

Using Administrative Data for
Clinical and Health Services Research

SID Data

- Florida SID Q4 2015–2019
 - Transition to ICD-10 in October 2015
- Records are all admissions from acute care community hospitals
 - Year (YEAR) and quarter (DQTR) refer to discharge
 - All-payer
- Florida hospitals, but patients not necessarily Florida residents
- Patients may have multiple records, i.e., multiple admissions
 - Duplicates are rare but do occur
 - Most common are interim discharge records, with same patient and admit date but different information

Data Files

- **Core** files contain most admission-related information
- **Charges** files record services provided that aren't captured by diagnosis/procedure codes
 - Correspond to revenue center codes
- **AHA Linkage** files associate hospital identifiers with those in the annual American Hospital Association survey databases
- HCUP **Cost-to-Charge Ratio** for Inpatient Files contain multipliers for converting from *charges* to *cost estimates*
 - **HTYPE** codes hospital ownership, size, and urban/rural status

Identifiers

- Each record in Core and Charges has unique `KEY`
- Each patient has unique `VisitLink`
 - Upon entering data, random number `DaysToEvent` assigned representing admission date
 - Indicates timing of subsequent events, e.g., admission 120 days later will have `DaysToEvent + 120`
 - `VisitLink` and `DaysToEvent` are retained across years
 - Quarter `DQTR` is as specific as we can be regarding calendar date
- In Core files, hospital is identified by `DSHOSPID`
 - AHA Linkage files can be used to get `HOSPID` (for cost-to-charge ratio files) and `AHAID` (for AHA survey files)

Core File Variables

- Over 200 variables
- Admission type (ATYPE) and disposition at discharge (DISPUNIFORM)
 - Hour of admission (AHOUR) and indicator for weekend admission (A WEEKEND)
- Expected primary payer (PAY1)
- Demographics:
 - AGE
 - FEMALE
 - RACE, HISPANIC
 - PL_CBSA, PL_NCHS, PL_RUCC, PL_UIC, PL_UR_CAT4
 - County and zip code
- Socioeconomic status:
 - Homeless
 - MEDINCSTQ, ZIPINC_QRTL

Diagnosis Codes

- *International Classification of Diseases, Tenth Revision, Clinical Modification* (ICD-10-CM) codes
 - Recorded in `I10_DXn` fields, between 31 and 34 per year

E Codes

- External cause of morbidity codes (ICD-10: V00–Y99)
 - Recorded in `I10_ECAUSEn` fields prior to 2017
 - Placed with other codes in `I10_DXn` fields starting in 2017

Present on Admission (POA) Codes

- `DXPOAn` and `E_POAn` fields
 - See documentation for meaning of different values

Procedure Codes

- *International Classification of Diseases, Tenth Revision, Procedure Coding System* (ICD-10-PCS) codes
 - Recorded in `I10_PRn` fields, 31 per year
 - No CPT/HCPCS codes

Timing Procedures

- `PRDAYn` fields record when procedures took place with respect to admission
 - 0 indicates day of admission, 1 indicates day after, ...
 - Negative values indicate days prior to admission

Resources

- SID database documentation
 - <https://www.hcup-us.ahrq.gov/db/state/sidbdbdocumentation.jsp>
- SID data elements by year (i.e., data dictionary)
 - https://www.hcup-us.ahrq.gov/db/state/siddist/siddist_ddeavailbyyear.jsp
- SID cost-to-charge ratio for inpatient files documentation
 - <https://www.hcup-us.ahrq.gov/db/ccr/ip-ccr/ip-ccr.jsp>
- List of variables in 1% SID data set with SAS labels
 - [admin_course_data/resources/variables_1p_sample.xlsx](#)
- List of variables in full SID data set with SAS labels
 - [admin_course_data/resources/variables_SID.xlsx](#)
 - [admin_course_data/resources/variables_comparison_by_year_SID.xlsx](#)