

## CSPH Data Set List

Please contact [csphdata@partners.org](mailto:csphdata@partners.org) if you would like to complete a project using a data set that is not currently available / not listed below so that we can discuss purchasing the data of interest. Data sets are purchased as needed following purchase requests. The requester must have funding available to support the purchase.

To request use of any of these data sets, please complete a [CSPH Data Core REDCap Request Form](#) to begin the process. A member of our team will then reach out to discuss next steps. Researchers must meet all IRB and Data Use Agreement requirements (database specific) before the Data Core will provide access.

### Data Sets Directly Available through the CSPH Data Core:

#### Healthcare Cost and Utilization Project (HCUP) – National Databases

- National Inpatient Sample (NIS)
  - 1988-2022
- Kids' Inpatient Database (KID)
  - 1997, 2000, 2003, 2006, 2009, 2012, 2016, 2019
- Nationwide Emergency Department Sample (NEDS)
  - 2006-2019, 2022
- Nationwide Readmissions Database (NRD)
  - 2010-2019
- Nationwide Ambulatory Surgery Sample (NASS)
  - 2016-2022

#### Healthcare Cost and Utilization Project (HCUP) – State Databases

*American Hospital Association Linkage files are available for most states and years (where made available by the state) and enable users to link the HCUP SID, SASD, or SEDD to the American Hospital Association survey data. NIS and KID files for years prior to 2011 can also be linked with the AHA survey.*

*California state data must be re-purchased for each individual project and is not available to share across CSPH.*

- State Inpatient Database (SID)
  - Arizona: 2009-2015, 2018-2021
  - Arkansas: 2010, 2014, 2021

- California: 2007-2011
- Colorado: 2008-2015
- District of Columbia: 2014
- Florida: 2008-2021
- Georgia: 2014
- Hawaii: 2014
- Iowa: 2014, 2021
- Kansas: 2014
- Kentucky: 2008-2016, 2018-2021
- Maryland: 2000-2019, 2021
- Massachusetts: 2008-2016, 2019
- Maine: 2014
- Michigan: 2014
- Minnesota: 2014, 2018-2021
- Mississippi: 2014, 2018-2021
- Nebraska: 2014, 2018-2021
- Nevada: 2014, 2016
- New Jersey: 2008-2022
- New Mexico: 2014
- New York: 2008-2016, 2018-2021
- North Carolina: 2008-2016, 2018-2021
- Oregon: 2014, 2018-2021
- Rhode Island: 2014, 2021
- South Carolina: 2014
- South Dakota: 2014
- Utah: 2014, 2018-2021
- Washington: 2008-2015 , 2018-2021
- West Virginia: 2014
- Wisconsin: 2014
  
- State Emergency Department Database (SEDD)
  - Arizona: 2021
  - Arkansas: 2014, 2021
  - California: 2005, 2011
  - Florida: 2008-2017, 2021
  - Iowa: 2021
  - Kentucky: 2014, 2016, 2021
  - Maryland: 2005, 2010-2017, 2021
  - Massachusetts: 2012, 2014, 2016, 2019
  - Minnesota: 2014, 2021
  - Nebraska: 2021
  - Nevada: 2014, 2016

- New Jersey: 2010-2022
- New York: 2008-2016, 2021
- North Carolina: 2016, 2021
- Oregon: 2021
- Rhode Island: 2021
- Utah: 2021
  
- State Ambulatory Surgery and Services Database (SASD)
  - California: 2007-2009, 2011, 2019, 2020
  - Florida: 2009-2014, 2016-2018
  - Kentucky: 2016
  - Maryland: 2009-2012, 2016
  - Nevada: 2016
  - New Jersey: 2010-2022
  - New York: 2009, 2012-2014, 2016
  - North Carolina: 2016

#### American Hospital Association (AHA) Annual Survey Database

- 2014, 2016, 2019, 2021-2023

#### National Emergency Medical Services Information System (NEMSIS)

- 2016 NEMSIS Public-Release Research Dataset

#### Trauma Quality Program (TQP)

- TQP combines data from the Trauma Quality Improvement Program (TQIP) and National Trauma Data Bank (NTDB). Records may contain data from either program, or both.
- 2007-2019 TQP Participant Use Files

#### Trauma Information Exchange Program (TIEP)

- 2014-2017; 2019-2022
- TIEP collects and maintains information regarding trauma centers in the U.S.
- Includes AHA hospital identification numbers to support linkage with HCUP, Medicare

#### CSPH Data Sets Shared at Principal Investigator's Discretion:

*These data sets may be available to CSPH researchers if the research fits within the project's Data Use Agreement. Interested researchers will need to receive permission from the Principal Investigator and may need to meet additional requirements to publish with the data. Additional fees may apply (ex: \$2,000 Medicare re-use fee for CMS).*

#### Medicare

- Medicare Serious Illness Sample (2016-2018, all fee-for-service beneficiaries with surgical care or selected serious illness diagnosis).

- Available for use with approval from purchasing PI (Joel Weissman)
- No fee for studies aligned with original DUA, \$2000 reuse fee (paid to CMS) for studies requiring a new DUA.
- Statistician fees may apply
- Contact [CSPHdata@partners.org](mailto:CSPHdata@partners.org) for more information

#### Tricare (administrative dataset)

- 2006-2022
- TRICARE is the healthcare insurance program for uniformed service personnel, their dependents, retirees, and retiree beneficiaries in the United States. Patients enrolled in TRICARE receive comprehensive medical coverage through either the Direct Care System (military facilities) or purchased care (civilian hospitals). The dataset includes administrative (i.e., billing) records for the entirety of their care.
- Inpatient and outpatient data sets available
  - Pharmacy
  - lab, radiology (no clinical results; just know tests were ordered)
- Dataset is administrative and includes:
  - Active on-duty (20%-ish),
  - reserve (small %),
  - dependents (50%-ish),
  - retirees (30%-ish)
- Statistician fees may apply
- Contact [CSPHdata@partners.org](mailto:CSPHdata@partners.org)

#### Additional Datasets Used by CSPH Researchers, but Not Provided by the Data Core

*These datasets are commonly used in projects at CSPH. However, the Data Core cannot provide you with access to an existing copy of the data set. Instead, we can support data storage and software resources, but investigators will need to independently request copies of the data from the original source.*

#### Other De-identified data sets available for free online

- National Health Interview Survey (NHIS)
- National Health and Nutrition Examination Survey (NHANES)
- Surveillance, Epidemiology, and End Results Program (SEER)
- National Health and Aging Trends Study (NHATS)
  - Some data elements are publicly available for free. Others require a more extensive request process if requesting a Limited Data Set.

#### SEER-Medicare

- Project specific and not available for outside project requests.

### National Cancer Database (NCDB)

- Request a letter of support through the CSPH REDCap form linked above. Requests and letters will be reviewed by the BWH/DFCI Committee on Cancer Chair.

### National Surgical Quality Improvement Program (NSQIP)

- NSQIP projects must be approved by [Jill Steinberg](#) and NSQIP on a project-by-project basis. CSPH cannot provide you with NSQIP data, although once approved by NSQIP you may use our server for data storage and analysis.