



Successfully Completing a National Center for Health Statistics Research Data Center Request for National Hospital Care Survey Data

Geoff Jackson M.S.

Division of Health Care Statistics

October 24, 2019

Agenda

- National Center for Health Statistics Research Data Centers
- Restricted National Hospital Care Survey data
- Requesting access to National Hospital Care Survey restricted-use data

National Center for Health Statistics Research Data Centers (RDC)

Research Data Center (RDC)

- The RDC is operated by the National Center for Health Statistics and allows researchers access to restricted-use data.
- The RDC is responsible for protecting the confidentiality of survey respondents, study subjects, or institutions while providing access to restricted-use data for statistical purposes.
- To access the data in the RDC you must submit a proposal outlining the requested analysis and results.
- The estimated project cost is \$3,000 with a range of \$1,800 to \$12,000. (For more info on fees: <https://www.cdc.gov/rdc/b5aprovproj/ap540.htm>)
- Complete the confidentiality training and forms. (<https://www.cdc.gov/rdc/b4confidisc/cfd400.htm>)

RDC Locations

- The NCHS RDC locations are:
 - Hyattsville, MD at NCHS headquarters
 - Atlanta, GA at the CDC Century Center Campus
 - Washington, DC at the Department of Health and Human Services headquarters
 - Rockville, MD at the Agency for Healthcare Research and Quality
- Federal Statistical RDC (formerly known as Census RDC) are located throughout the country and are partnerships between federal statistical agencies and leading research intuitions. *Note: FSRDCs may charge additional fees.*

FSRDC locations



Kentucky FSRDC

- Located in the Gatton College of Business and Economics on the University of Kentucky campus.
- Collaboration between the University of Kentucky and the U.S. Census Bureau, and includes consortium partners of Indiana University, The Ohio State University, University of Cincinnati, and University of Louisville.
- Includes data from the Census Bureau, Bureau of Labor Statistics, and the National Center for Health Statistics.



National Hospital Care Survey (NHCS) Data in the RDC

National Hospital Care Survey (NHCS)

Goal and Objectives

- Goal:
 - Provide reliable and timely healthcare utilization data for hospital-based settings.
- Objectives:
 - Move toward electronic data collection, particularly electronic health records.
 - Provide benchmark data for comparison to national data.
 - Link episodes of care across hospital units as well as link to other data sources such as the National Death Index (NDI) and CMS Medicare data.

NHCS Sample Design, Data Coverage, and Sources

- Hospitals are randomly selected to provide nationally representative data on hospital utilization. The sampled hospitals represent facilities of similar size, service type, and/or geographic location and cannot be replaced.
- The 2019 NHCS sample consists of 598 non-institutional, non-federal hospitals with six or more staffed inpatient beds.
- Participating hospitals are asked to submit all inpatient discharges and emergency department (ED) visits for up to a 12-month period. Outpatient department (OPD) data was last collected in 2016 and may be added in the future.
- Data Sources: UB-04 administrative claims, electronic health records (EHR), and Vizient

NHCS Data in the RDC

- There are approximately **125 million** unique hospital encounters available from 2013 through 2016.
- Encounters are collected from the Inpatient Department (IP), Emergency Department (ED), and Outpatient Department (OPD).

Table 1. Number of Hospitals and Encounters by Department by Year

Year	IP		ED		OPD		Total	
	Hosp	Encounters	Hosp	Encounters	Hosp	Encounters	Hosp	Encounters
2013	97	1,474,478	82	3,784,397	87	15,144,448	97	20,403,323
2014	94	1,653,622	83	4,530,360	86	19,005,777	95	25,189,759
2015	114	2,204,258	97	5,900,738	101	26,455,149	118	34,560,145
2016	145	2,591,722	124	7,032,304	128	35,692,420	150	44,572,315

Type of NHCS Data Available in the RDC

- Core NHCS data elements:
 - Survey
 - Patient
 - Encounter
- Linked data
 - National Death Index (NDI)
 - NDI supplemented with results from the Drug Involved in Mortality (DIM) file
 - CMS Master Beneficiary Summary File

NHCS Core Data Elements

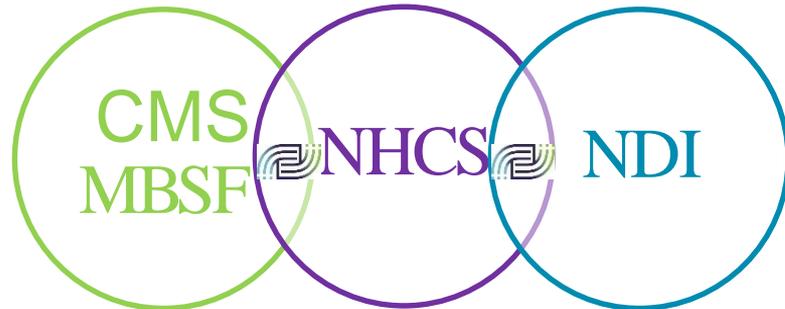
- <https://www.cdc.gov/rdc/b1datatype/Dt1224h.htm>

The screenshot displays the CDC National Center for Health Statistics website. At the top left is the CDC logo with the text "Centers for Disease Control and Prevention" and "CDC 24/7. Saving Lives. Protecting People™". To the right is a search bar with the text "Search" and a magnifying glass icon. Below the search bar is a blue navigation bar with the text "National Center for Health Statistics". Underneath this bar is a breadcrumb trail: "CDC > NCHS > RDC > Restricted Data > NCHS Restricted". To the right of the breadcrumb trail are social media icons for Facebook, Twitter, LinkedIn, YouTube, and Instagram. On the left side of the page is a sidebar menu with a "RDC" header and a "Restricted Data" section. Under "Restricted Data" are several sub-sections: "NCHS Public Use", "NCHS Restricted" (which is currently selected), "NHANES Restricted Variables", "NAMCS and NHAMCS Restricted Variables", "NHDS Restricted Variables", "NNHS and NNAS Restricted Variables", "NHHCs and NHHS Restricted Variables", and "NSRCF Restricted Variables". The main content area features the title "National Hospital Care Survey (NHCS)" in a large, bold font. Below the title is a paragraph describing the survey: "The [National Hospital Care Survey](#) (NHCS) is designed to provide accurate and reliable health care statistics that answer key questions of interest to health care and public health professionals, researchers, and health care policy makers." This is followed by another paragraph: "The survey describes national patterns of health care delivery in hospital-based settings, including inpatient departments, emergency departments (EDs), and outpatient departments (OPDs). Data elements collected include demographics, diagnoses, procedures, laboratory tests, and medications. Additionally, NHCS collects patient-level identifiers that allow a patient's episodes of care to be linked between hospital inpatient and outpatient settings, as well as to outside databases such as the National Death Index, providing a more complete picture of patient care." Below this is a third paragraph: "The 2013-2016 unweighted inpatient and ambulatory National Hospital Care Survey datasets are now accessible through the NCHS Research Data Center (RDC)." At the bottom of the main content area are four links, each with a PDF icon and file size: "[2013-2014 NHCS Data Dictionary](#). [PDF - 114 KB]", "[2015 NHCS Data Dictionary](#). [PDF - 362 KB]", "[2016 NHCS Data Dictionary](#). [PDF - 1 MB]", and "[NHCS Documentation and User's Guide](#). [PDF - 391 KB]".

External Data Linked to the NHCS Data

- **NDI** is a centralized database of death record information on file in state vital statistics offices; and
- **MBSF** includes information on beneficiary demographic characteristics, reason for Medicare entitlement, and program enrollment type (Original Medicare vs. Medicare Advantage).

https://www.cdc.gov/nchs/data-linkage/nhcs_linkage.htm



Data Source: National Death Index (NDI)

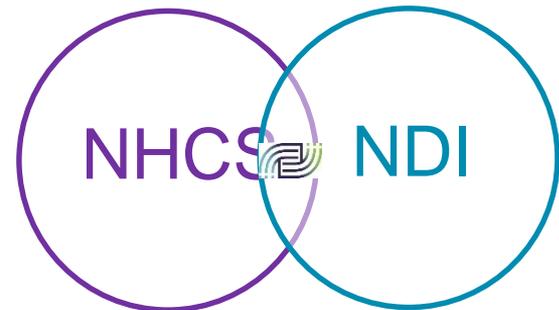
- Centralized database of death record information on file from state vital statistics offices in the U.S.
 - Housed at National Center for Health Statistics
- Death records added annually from 1979



<https://www.cdc.gov/nchs/ndi/index.htm>

NHCS – NDI Linkage

- Linked NHCS-NDI data offer opportunities to study mortality post hospital discharge
- 2014 NHCS data are linked to the 2014 and 2015 NDI file.
- 2016 NHCS data are linked to the 2016 and 2017 NDI file.
- The 2014 and 2016 NHCS data are not nationally representative.
 - Sample weights are not available.
 - Unweighted estimates may have bias.



NHCS – NDI Data Available in the RDC

- <https://www.cdc.gov/nchs/data-linkage/nhcs-ndi.htm>

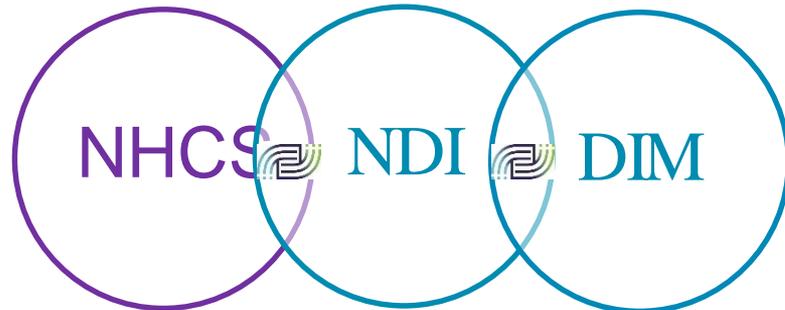
NHCS Linked to the NDI Mortality Variables
 Date Created: 12AUG2019
 Number of Variables: 65

Variable Name	Variable (VAR) Label	VAR Type	Range of Values	Value Description	Notes ^{1, 2}
SURVEY	Survey Name	Char			The values for this variable are: 'NHCS 2014' or 'NHCS 2016'
PATIENT_ID	Patient Identification Number	Char	ID		Public use identifier assigned by NCHS. Researchers linking to the NHCS analytic file should use PATIENT_ID as the common key.
ELGSTAT	Linkage Eligibility Status	Num	0 1	0: Ineligible 1: Eligible	Survey participants are defined as ineligible for linkage if they had insufficient identifying data to create a NHCS submission record. Please note that all survey participants are included on the linked NHCS files regardless of linkage eligibility.
MORTSTAT	Final Mortality Status	Num	0 1 2 3 Missing	0: Eligible for data linkage, assumed alive 1: Eligible for data linkage, assumed deceased based on NDI linkage 2: Eligible for data linkage, assumed deceased from non-NDI source 3: Ineligible for data linkage, assumed deceased from non-NDI source Missing: Ineligible for data linkage, no other source of death available	Blank values imply linkage ineligibility. The MORTSTAT variable is NCHS's final determination of vital status and should be used as an outcome variable and to calculate survival. Codes are assigned depending on eligibility and vital status.
MORTSRCE_NDI	Mortality Source: NDI Match	Num	1 Missing	1: Mortality status ascertained through a probabilistic match to a NDI record Assumed alive, ineligible for linkage, or death by source other than NDI linkage.	MORTSRCE_NDI=1 indicates that mortality status was ascertained through a probabilistic match to a National Death Index (NDI) record.

¹2014 Mortality Multiple Cause-of-Death Record Layout: https://www.cdc.gov/nchs/data/ndi/Record_Layout_2014.pdf (accessed on August 06, 2019)
²2016 Mortality Multiple Cause-of-Death Record Layout: https://www.cdc.gov/nchs/data/ndi/Multiple_Cause_Record_Layout_2016.pdf (accessed on August 06, 2019)

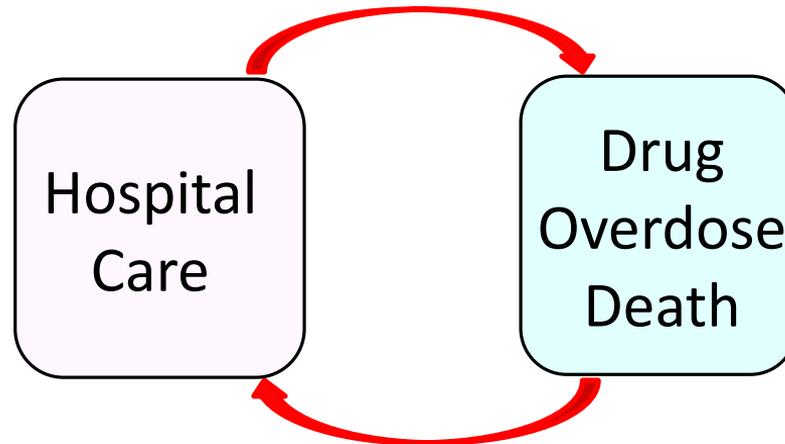
Data Source: Drug Involved in Mortality

- Drug Involved in Mortality (DIM), formerly known as NVSS-M-DO, data provide information on specific drugs involved in deaths, extracted from the literal text of death certificates.
- The methodology for constructing the DIM data was developed collaboratively between the National Center for Health Statistics and the Food and Drug Administration.



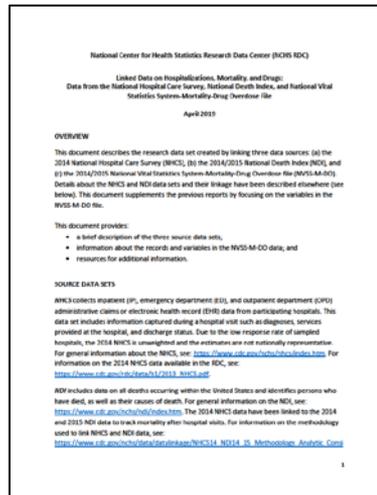
NHCS – NDI – DIM Linkage

- Linked NHCS – NDI – DIM data allow researchers to study specific drugs, drug classes, and non-specific references to drugs mentioned in the literal text of the death certificate.
- Allows for look forward and look back analyses.



NHCS – NDI – DIM Data Available in the RDC

- 2014 NHCS linked to the 2014 and 2015 NDI and DIM
- **Available in 2020:** 2016 NHCS linked to the 2016 and 2017 NDI and DIM
- https://www.cdc.gov/rdc/data/b1/PCOR_III_data_doc.pdf



Data Source: CMS Master Beneficiary Summary File

- Contains data on all Medicare beneficiaries enrolled in/or entitled to Medicare within the calendar year
- Includes reason for entitlement
 - Medicare and Medicaid dual eligibility status
- Indicates participation in Medicare Advantage or Fee-for-Service Medicare
- Allows for leveraging information on comorbidities (e.g., CMS' Chronic Conditions Data Warehouse) and Medicare coverage that may be associated with hospital-related service outcomes and health outcomes

NHCS – CMS Linkage

- **Base (A/B):** beneficiary characteristics, monthly entitlement indicators, reasons for entitlement (initial and current), and monthly Medicare Advantage indicators.
- **Part D:** variables specific to Medicare Part D Prescription Drug Plan.
- **Cost & Utilization:** summarized information about the service utilization and Medicare payment amounts by type of claim, including prescription drugs.
- **Conditions:** variables that indicate a Medicare beneficiary has received a service or treatment for selected chronic health conditions.



Requesting Access to the NHCS RDC Data

Requesting Access with a Research Proposal

- Researchers must submit a research proposal.
- The proposal must have the following criteria:
 - A well-defined research question that addresses a public health concern.
 - Explanation of what restricted-use variables are needed to complete the project and why.
 - The disclosure risk associated with:
 - The requested restricted-use variables
 - The requested mode of access
 - Analytic plan (this includes statistical methods) and
 - The nature and composition of your planned output.

RDC Analyst's Role

- Each proposal will be assigned an RDC Analyst.
 - Facilitates the review
 - Creates the analytic data set and provides it to the RDC location described in your proposal
 - Accepts payment
 - Accepts NCHS Confidentiality required paperwork
 - Reviews your output for disclosure risk
 - Provides your approved output

Proposal Process

1. Determine the need and identify the restricted-use data.
2. Determine preferred location of access.
3. Draft your research proposal.
4. Submit your proposal (using the Proposal Format, include page numbers) as one document to rdca@cdc.gov
5. Wait for comments from the NCHS Review Committee and respond quickly to expedite any proposal revisions.
6. Updated your proposal when there are changes.

Drafting a Proposal

- The research proposal must comply with the format in the RDC proposal document.
- <https://www.cdc.gov/rdc/data/b3/Proposal-Format-v3.2.pdf>

RDC Research Proposal	
Proposal Format Updated 09/2019	
General Information	
Name:	
Title of Project:	
RDC Data System and Years:	
Non-RDC State File:	
State of Access:	<input type="checkbox"/> NCHS RDC, Hyattsville, MD <input type="checkbox"/> NCHS RDC, Rockville, MD <input type="checkbox"/> NCHS RDC, Washington, DC (Government Only) <input type="checkbox"/> NCHS RDC, Atlanta, GA <input type="checkbox"/> External Storage/RDC specific
Statistical Software:	<input type="checkbox"/> SAS/SDAAN <input type="checkbox"/> State <input type="checkbox"/> Other, specify _____
(Does not have special):	
Proposed Start Date:	
Funding Source:	
Billing Address:	
(Include contact person)	

Common Mistakes to Avoid

- Do not attempt to misuse the data.
 - NHCS data are not nationally representative. It cannot be used to produce estimates at the state or national level.
- Do not request results that could be used to identify a hospital or patient.
 - Asking for information produced at the county level or zip code level could be used to identify the participating hospital.
 - Producing results on rare conditions by low geographic levels or month/date of birth.
- Reduce the need for amendments:
 - Identify all intended results prior to submission and include the tables in the proposal.
 - Identify all necessary variables in the submission.

Merging Non-NCHS Data in the RDC

- Non-NCHS data can be merged to restricted use NCHS data in the RDC.
- Researcher must provide the data and identify how to link the data to the data in the RDC.
- Example: Identify urban/rural designation of patients in the NHCS data
 - Patient residence Zip code is available in the NHCS data.
 - Provide a file with urban/rural designation assigned to Zip code.
 - Include a data dictionary of the non-NCHS data in the proposal.
 - Detail the merging of the urban/rural information by Zip code in the proposal.
 - Detail the intended output by urban/rural designation in the proposal.

Proposal Review

- The proposal is reviewed by a committee.
- The proposal will not be reviewed for scientific merit.
- There are three outcomes of the review:
 - Revise and resubmit
 - Approve
 - Disapprove

2019 Call for Applications

NCHS/AcademyHealth Health Policy Fellowship Program

Conduct Research Using National Center for Health Statistics Restricted Data Files

PURPOSE

This Fellowship Program promotes a wide variety of health services research that uses data from the National Center for Health Statistics (NCHS). The program also offers collective opportunities with AcademyHealth and NCHS.

AS A FELLOW, YOU WILL:

- Conduct research for up to 12 months at NCHS using restricted data files with strong preference for data linked to external sources;
- Attend and present research at the AcademyHealth Annual Research Meeting; and
- Network with current and past AcademyHealth Fellows and Scholars.

For more information on the NCHS Data Linkage Program, visit <https://www.cdc.gov/nchs/data-linkage/index.htm>.
The application cycle is slated to open in October 2019.

To be informed when the application cycle opens, or if you have any questions, please contact nchs@academyhealth.org.



Contact Information

- **Geoff Jackson**

email: mlq2@cdc.gov

phone number: 301-458-4703