RDC Health Data
Data available from the
MU Research Data Center
A Branch of the Kansas City RDC and a unit of the
Federal Statistical Research Data Centers (FSRDC)

Jake Cronin, MU RDC Administrator (Presenter)
Eileen Avery, Director
May 20, 2020
Presentation Outline

• Background and purpose of FSRDCs
• Overview of health data available through RDC
  – National Center for Health Statistics (NCHS)
  – Agency for Healthcare and Research Quality
• Process for accessing RDC
• Contact information and sources for additional information
• Questions
What is a Federal Statistical Research Data Center?

- A secure environment where researchers can undertake analysis of non-public data from federal agencies
  - Located in Ellis Library

- MU joined around 30 RDC sites nationwide.

- The RDC allows for research that would be difficult (if not impossible) otherwise
National Center for Health Statistics (NCHS)

This section will review a suite of surveys available through NCHS as well as restricted use data only available in the RDC.
Types of Data and Data Access

• **Public**
  • Data files: collapsed categories, top and bottom coded
  • Aggregated reports or tabulations

• **Restricted: Direct Identifiers**
  • Name, SSN, address
  • Used to create the Linkage Products and removed
  • *Never available*

• **Restricted: Indirect Identifiers**
  • Items that when combined with other information (from NCHS or other sources) could lead to identification
  • *Available through the RDC*
National Health Interview Survey

• In person, in home survey collects data on about 100,000 persons in 40,000 responding households per year representative of the non-institutionalized civilian U.S. population

• Data on:
  • Health status, health conditions, functioning and disability
  • Insurance coverage
  • Access to and use of health services
  • Immunization
  • Health-related behaviors and risk factors
  • Special topics
Publications using restricted-use NHIS data


National Health and Nutrition Examination Survey

• Standardized physical examinations, laboratory tests, personal interviews of around 5,000 persons representative of the non-institutionalized civilian U.S. population
• Data on:
  • Disease/condition prevalence
  • Nutrition
  • Body measures
  • Growth and development
  • Health-related behaviors
  • Risk factors
Publications using restricted-use NHANES data

National Survey of Family Growth

• In-person interview with men and women 15-44 years of age, representative of the non-institutionalized civilian U.S. population

• Data on:
  • Reproductive health
    • Fertility/infertility
    • Contraception
    • Pregnancy
    • Sexual activity
  • Family formation
    • Marriage, divorce, cohabitation
Publications using restricted-use National Survey of Family Growth data


State and Local Area Integrated Telephone Survey

  • National Survey of Adoptive Parents (2007)
  • National Survey of Children in Nonparental Care (2013)
  • National Survey of Adoptive Parents of Children with Special Health Care Needs (2008)
  • Survey of Pathways to Diagnosis
Publications using restricted-use State and local area integrated telephone survey data


National Immunization Survey

• Telephone interview of households and mail survey of doctors and other vaccination providers to monitor immunization coverage for children (NIS: 18-35 months) and adolescents (NIS-TEEN: 13-17 years).

• Data on:
  • Types of immunizations
  • Dates of administration
  • Facility characteristics of vaccination providers
Publications using restricted-use National Immunization Survey data

- Gilkey MB, Magnus BE, Reiter PL, et al. (2014). “The Vaccination Confidence Scale: A brief measure of parents' vaccination beliefs.” *Vaccine*
National Health Care Surveys

- Ambulatory and hospital care surveys
  - National Ambulatory Medical Care Survey (NAMCS)
    - Physician offices
    - Community Health Centers
  - National Hospital Ambulatory Medical Care Survey (NHAMCS)
    - Outpatient Clinics
    - Emergency rooms
  - National Hospital Care Survey (NHCS)
    - Inpatient care -- hospitals
National Health Care Surveys

• Long-term care surveys
  • National Survey of Residential Care Facilities (NSRCF)
  • National Home and Hospice Care Survey (NHHCS)
  • National Nursing Home Survey (NNHS)
  • National Study of Long-Term Care Providers (NSLTCP)
    • Bi-annual data collection starting in 2012
    • Administrative data: home health agencies, nursing homes, hospices
    • Questionnaire data: adult day services centers, residential care communities
Publications using restricted-use National Health Care Survey data

Linked data enable broader analyses of factors that influence health and health outcomes.

Surveys are linked with administrative data such as:
- The National Death Index
- Claims data from the Centers from Medicare & Medicaid Services
- Social Security Benefit History data

Linked data are accessed through:
- Public-use files
- NCHS Research Data Center (for restricted use files)
National Vital Statistics System

- Data received from 57 vital registration jurisdictions (50 states, DC, NYC, 5 territories)
  - Births and information on the pregnancy and delivery
  - Deaths and information on cause of death
- Produces state and local level data
NCHS Linked Data

Linked to Mortality
- NHIS 1985-2014
- NHANES: 1999-2014
- NNHS: 85, 95, 97, 04

Linked Social Security Benefit History
- NHIS 1994-2005
- NHANES, 99-04

Linked to Housing and Urban Development
- NHIS: 1999-2016
- NHANES: 1999-2016

Linked to Medicare
- NHIS 1994-2013
- NHANES: 1999-2012

Linked to Medicaid
- NHIS 1994-2013
- NHANES, 1999-2012
- NNHS 2004

End Stage Renal Disease Files
- NHIS: 1994-2005
- NHANES: 1999-2004
Publications using restricted-use linked data

• Nachman KE, Parker JD. (2012). “Exposures to fine particulate air pollution and respiratory outcomes in adults using two national datasets: a cross-sectional study.” *Environmental Health*.

Substance Abuse and Mental Health Services Administration (SAMHSA)

- National Survey on Drug Use and Health (NSDUH)
  - use of illegal drugs, prescription drugs, alcohol, and tobacco
  - mental disorders, treatment, and co-occurring substance use and mental disorders

- NSDUH data provides estimates of substance use and mental illness at the national, state, and sub-state levels. NSDUH data also helps to identify the extent of substance use and mental illness among different sub-groups, can be used for estimates of trends over time, and measures the need for treatment services
United States Cancer Statistics

- The United States Cancer Statistics (USCS) are the official government statistics on cancer. The files are comprised of data supported by both the Centers for Disease Control and Prevention and the National Cancer Institute.
- More than 24 million observations.
- Data includes demographic variables such as age, sex, and race as well as tumor characteristics (for example, year of diagnosis, site, histology, stage, and behavior).
- Restricted data includes things such as: state and county identifiers, detailed race categories, detailed diagnostic categories, exact age at diagnosis, and others.
Agency for Healthcare and Research Quality (AHRQ)

This section will review a suite of surveys available through AHRQ as well as restricted use data only available in the RDC.
Medical Expenditure Panel Survey (MEPS)

- MEPS collects data on the specific health services that Americans use, how frequently they use them, the cost of these services, and how they are paid for, as well as data on the cost, scope, and breadth of health insurance held by and available to U.S. workers.
- Collects data from families and individuals, their medical providers (doctors, hospitals, pharmacies, etc.), and employers across the United States.
- Panels consist of several rounds of interviewing covering two full calendar years.
  - Makes it possible to determine how changes in respondents' health status, income, employment, eligibility for public and private insurance coverage, use of services, and payment for care are related.
Household Component:

- Nationally representative sample of families and individuals (subsample of households that participated in NHIS)
- Collects detailed information for each person in the household on the following: demographic characteristics, health conditions, health status, use of medical services, charges and source of payments, access to care, satisfaction with care, health insurance coverage, income, and employment.
- 13,936 families consisting of 35,068 people in 2013 sample
Insurance Component

- Collects data from a sample of private and public sector employers on the health insurance plans they offer their employees.
- The collected data include the number and types of private insurance plans offered (if any), premiums, contributions by employers and employees, eligibility requirements, benefits associated with these plans, and employer characteristics.
- Two distinct samples:
  - Nationally representative independently drawn, nationwide sample of establishments and state/local governments. Controlled by Census Bureau.
  - The MEPS-IC Household sample was a sample of employers that were identified by respondents in the MEPS-HC as their main employer or secondary employer that was the source of their health insurance. Controlled by AHRQ.
Confidential AHRQ Data

- Household Component–Insurance Component Linked File
- Nursing Home Component (available for 1996 only)
- Medical Provider Component
- Area Resource File
- MEPS Link Files to NHIS
Confidential and non-public use variables

- Fully Specified ICD-9 Codes: These codes allow medical conditions to be identified with greater specificity.
- Fully Specified Industry and Occupation Codes.
- State and County FIPS Codes
- Census Tract and Block-Group Codes
- Federal and State Marginal Tax Rates
  - Tax simulations using the National Bureau of Economic Research’s TAXSIM package have been run for the 1996–2002 HC full-year populations. Tax amounts and marginal tax rates have been computed for Federal, State and FICA taxes. Property taxes, sales taxes, and city/county income taxes have not been simulated.
MU RDC Health Data Projects

• A researcher has been working with the restricted-use NHIS to study the effects of the ACA Medicaid expansion on health and access to healthcare. Needed the restricted use for state identifiers.
• Several researchers have examined various immigrant groups. For example, a researcher in the Truman School looked at how county-level contextual factors predict immigrant’s health and well-being.
• Lastly, another researcher looked at how the ACA effected access to care for people with a very particular kidney disease. Needed restricted data for detailed ICD-9 codes.
COVID-19 Research

• Federal agencies are working to expedite the proposal process for researchers interested in studying topics related to Covid-19

• Census currently conducting the Household Pulse Survey
  – Collecting data on the socioeconomic impact of Covid-19 on American households. Survey was just conducted recently.

• Most data is dated but still possible to link current Covid data to past aggregated data to determine patterns and trends
  – Lisa Sattenspiel project
RDC Proposal Process
Overview

1. Researchers interested in using restricted-use datasets will need to go through a proposal process and obtain Special Sworn Status (SSS).
2. Jake Cronin is the Census RDC Administrator that works full-time at the MU RDC to help researchers through the proposal process.
3. For NCHS and AHRQ proposals, researchers work with an analyst at those respective agencies to draft a proposal. For simplicity, you can contact Jake and he can direct you to the appropriate agency.
Proposal Process

• Process can vary by agency, but generally:
  – Work with analyst at NCHS/AHRQ to draft proposal
  – Proposal is submitted to NCHS/AHRQ and is reviewed by Review Committee
  – NCHS/AHRQ emails researcher and MU RDC Administrator when proposal is approved
  – Pay data extract fee (this is the money paid to the agency to assemble your data)
  – Your approved data is uploaded to directory on Census servers and accessed at MU RDC
  – Still accepting proposals during shutdown!
Special Sworn Status

• After the proposal has been approved by the appropriate agency, researchers will need to obtain Special Sworn Status (SSS) in order to physically access the RDC.
  – SSS entails a background check and an oath to project the confidentiality of the data forever
• Requires completing paperwork, getting fingerprints, online trainings, and identification verifications.
• Process can take a few months to complete
• Not currently accepting SSS applications during shutdown, but there are plans to restart the process soon.
Working in an RDC

• Only researchers with approved projects and Special Sworn Status are allowed in the RDC.
• NCHS/AHRQ researchers can only work in the RDC when Administrator is present
• Researchers will work on “thin client” computers that tunnel into secure Census servers.
  – No data is stored on local machines
• All analysis must be done in the RDC.
• Output cannot be removed from the RDC prior to a formal disclosure avoidance review
Disclosure Avoidance Review

• There is a formal process to remove aggregated output out of the RDC.
• This is the only approved process for moving results from inside the confidential RDC to the outside public domain.
• The general idea is to ensure that no identifiable information on establishments, people, or households is disclosed.
  – In the proposal phase, it is important to think ahead about the output of the research and if it poses any disclosure risk
More Information

• National Center for Health Statistics (NCHS):
  Website: http://www.cdc.gov/nchs/r&d/rdc.htm
  Email: rdca@cdc.gov

• Agency for Healthcare Research and Quality (AHRQ)
  Website: http://www.meps.ahrq.gov/mepsweb/data_stats/onsite_datacenter.jsp
  Email CFACTDC@AHQR.HHS.GOV
  Ray.Kuntz@ahrq.hhs.gov
MU RDC Contact Information

Jacob Cronin
MU RDC Administrator
Jacob.Cronin@census.gov
573-884-9122

Eileen Avery
Director
AveryE@Missouri.edu