National Center for Health Statistics
Data Linkage

# The NCHS Data Linkage Program: Leveraging the nation's health data for evidence-based decision making

**Lisa B. Mirel**

**Director, Data Linkage Program**

**National Committee on Vital and Health Statistics**

**September 9, 2021**

# National Center for Health Statistics (NCHS)

- Nation’s principal health statistics agency
- One of 13 federal statistical agencies
- **Mission:** To provide statistical information that will guide actions and policies to improve the health of the American people. As the Nation’s principal health statistics agency, NCHS leads the way with accurate, relevant, and timely data.

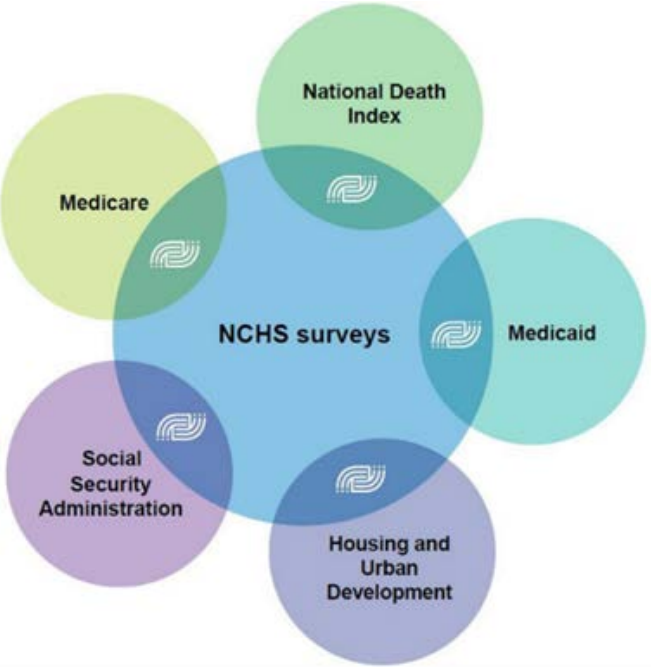
# Policy Questions Require Evidence-Based Answers

- Many pressing policy questions require complex, detailed data

  - *Do Social Security Disability Insurance beneficiaries have access to care during the waiting period before Medicare entitlement?*
  - *Are there adverse health effects associated with the mandatory folic acid fortification policy for grain products?*
  - *How effective are health and housing policies in reducing lead exposure?*

# Data Linkage as a Solution

- Linking data is a powerful mechanism to provide policy relevant information in an efficient way
  - Health survey data are collected to monitor health status, health behaviors and health care access
  - Administrative data are collected for programmatic purposes
- Combining these types of existing data efficiently creates opportunities to answer key health and policy relevant questions



# NCHS Surveys Included in Linkages

## National Health Interview Survey (NHIS)

A nationally representative, cross-sectional household interview survey that serves as an important source of information on the health of the civilian noninstitutionalized population of the United States

## National Health and Nutrition Examination Survey (NHANES)

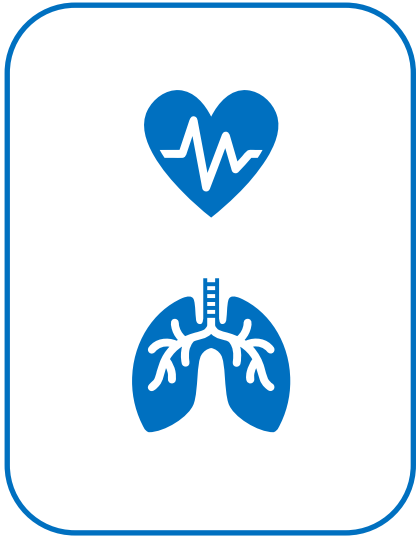
A nationally representative, cross-sectional sample of the U.S. civilian noninstitutionalized population, which includes an interview in the household followed by an examination in a mobile examination center

## National Health Care Surveys

A family of data collection efforts that gather information about providers of health care services and the patients they serve across the spectrum of health care settings from ambulatory and hospital care to long term care settings

# NCHS Surveys Collect Information on

Health behaviors



Health conditions

Socioeconomic status

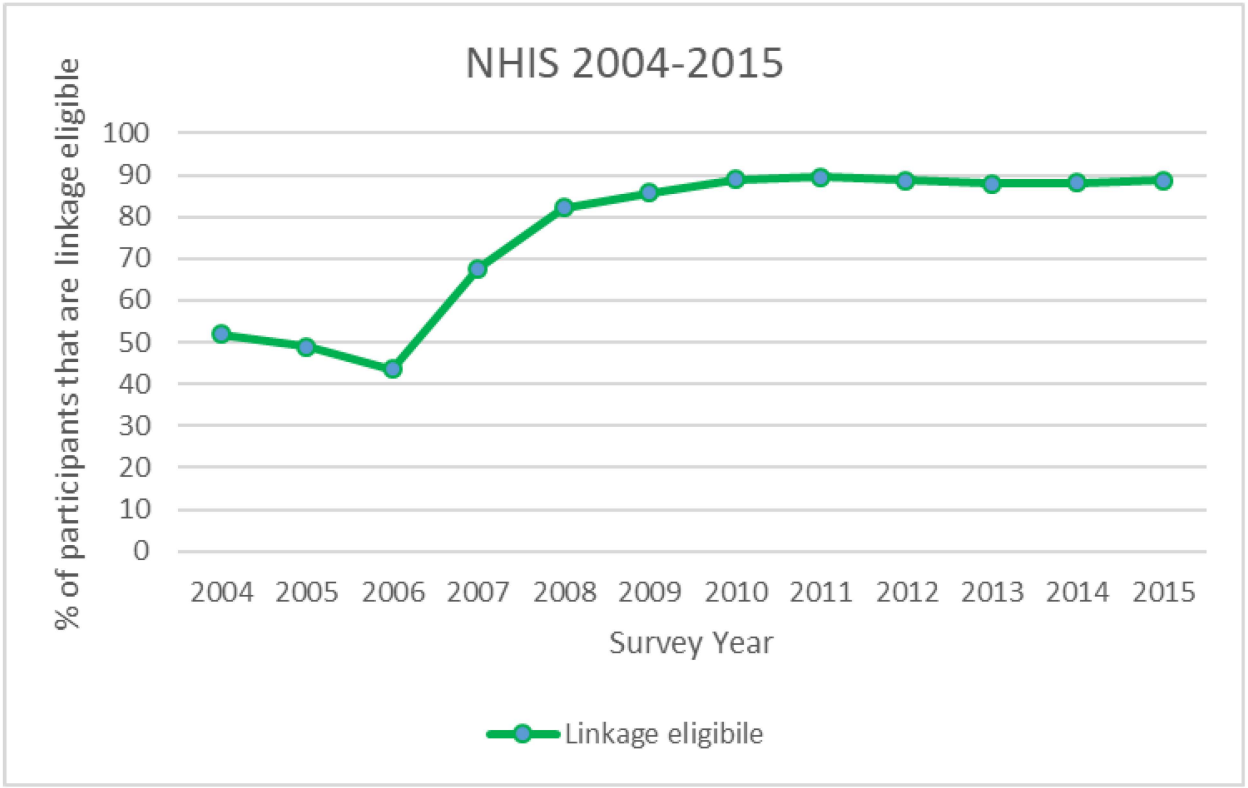
Healthcare access and utilization

# Linkage Eligibility

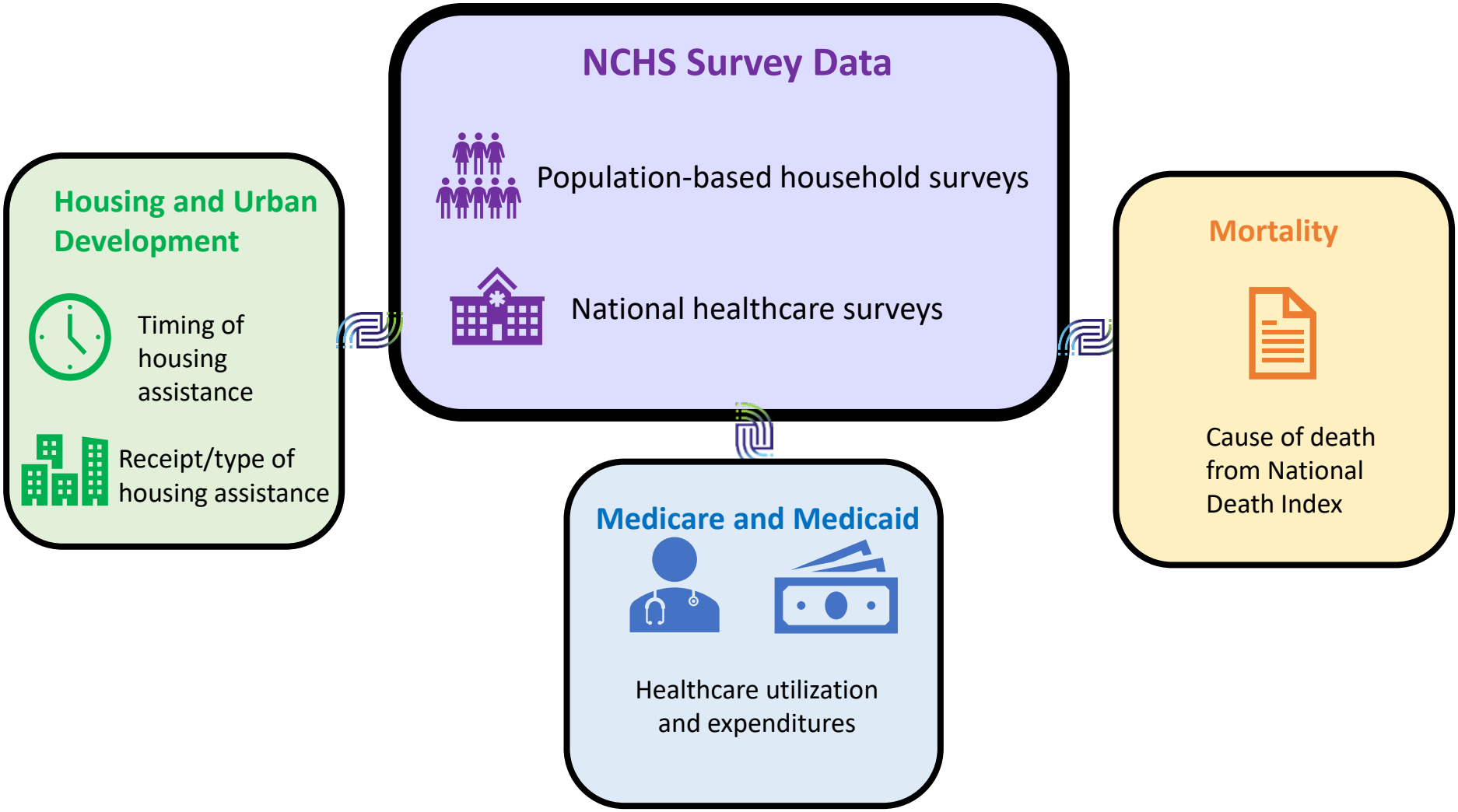
- Only NCHS survey participants who have provided consent as well as the necessary personally identifiable information (PII) are considered ***linkage-eligible***

- NCHS survey participants are informed of NCHS' intent to conduct data linkage activities through a variety of informed consent procedures:
  - Advance letters
  - Participant brochures
  - Signed consent forms
  - Questionnaires

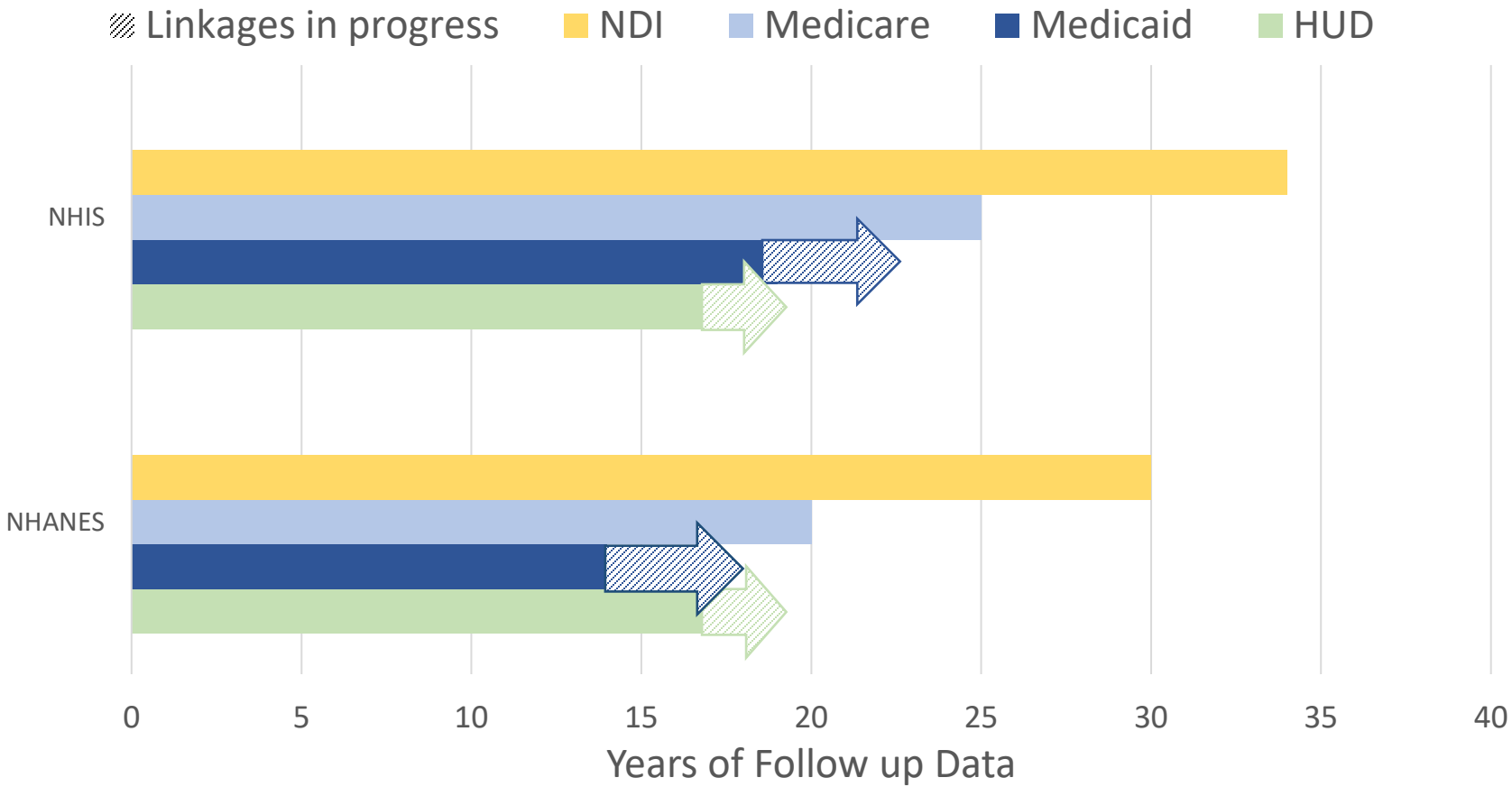
# Percent of NHIS Participants that are Linkage Eligible




Note: NHIS is the National Health Interview Survey; this graph only includes sample adults/ sample children


NCHS Survey Data
Population-based household surveys
National healthcare surveys
Housing and Urban Development
Timing of housing assistance
Receipt/type of housing assistance
Mortality
Cause of death from National Death Index
Medicare and Medicaid
Healthcare utilization and expenditures

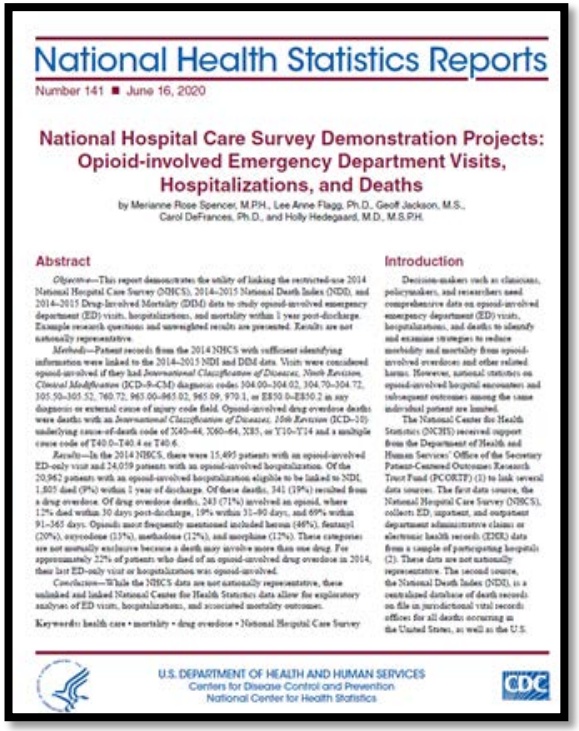
# Availability of Linked NHIS and NHANES Data



# Linked National Hospital Care Survey data

- Linked electronic health records and claims from the National Hospital Care Survey with Medicare data to provide new resources for patient outcomes research, including evaluating treatment effectiveness and post-acute care utilization

- Linked Inpatient and Emergency Department records with death certificate data to assess 30, 60, 90-day post hospitalization mortality



National Health Statistics Reports

Number 141 ■ June 16, 2020

**National Hospital Care Survey Demonstration Projects: Opioid-involved Emergency Department Visits, Hospitalizations, and Deaths**

by Merianne Rose Spencer, M.P.H., Lee Anne Flagg, Ph.D., Geoff Jackson, M.S., Carol DeFrances, Ph.D., and Holly Hedegaard, M.D., M.S.P.H.

**Abstract**

*Objectives*—This report demonstrates the utility of linking the restricted-use 2014 National Hospital Care Survey (NHCS), 2014–2015 National Death Index (NDI), and 2014–2015 Drug-Involved Mortality (DIM) data to study opioid-involved emergency department (ED) visits, hospitalizations, and mortality within 1 year post-discharge. Example research questions and unweighted results are presented. Results are not nationally representative.

*Methods*—Patient records from the 2014 NHCS with sufficient identifying information were linked to the 2014–2015 NDI and DIM data. Visits were considered opioid-involved if they had *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) diagnosis codes 304.00–304.02, 304.70–304.72, 305.50–305.52, 760.72, 965.00–965.02, 965.09, 970.1, or E850.0–E850.2 in any diagnostic or external cause of injury code field. Opioid-involved drug overdose deaths were deaths with an *International Classification of Diseases, 10th Revision* (ICD-10) underlying cause-of-death code of X40–44, X60–64, X85, or Y10–Y14 and a multiple cause code of T40.0–T40.4 or T40.6.

*Results*—In the 2014 NHCS, there were 15,495 patients with an opioid-involved ED-only visit and 24,059 patients with an opioid-involved hospitalization. Of the 20,962 patients with an opioid-involved hospitalization eligible to be linked to NDI, 1,805 died (9%) within 1 year of discharge. Of these deaths, 341 (19%) resulted from a drug overdose. Of drug overdose deaths, 243 (71%) involved an opioid, where 12% died within 30 days post-discharge, 19% within 31–90 days, and 69% within 91–365 days. Opioids most frequently mentioned included heroin (46%), fentanyl (20%), oxycodone (15%), methadone (12%), and morphine (12%). These categories are not mutually exclusive because a death may involve more than one drug. For approximately 22% of patients who died of an opioid-involved drug overdose in 2014, their last ED-only visit or hospitalization was opioid-involved.

*Conclusion*—While the NHCS data are not nationally representative, these unlinked and linked National Center for Health Statistics data allow for exploratory analyses of ED visits, hospitalizations, and associated mortality outcomes.

Keywords: health care • mortality • drug overdose • National Hospital Care Survey

**Introduction**

Decision-makers such as clinicians, policymakers, and researchers need comprehensive data on opioid-involved emergency department (ED) visits, hospitalizations, and deaths to identify and examine strategies to reduce morbidity and mortality from opioid-involved overdoses and other related harms. However, national estimates on opioid-involved hospital encounters and subsequent outcomes among the same individual patient are limited.

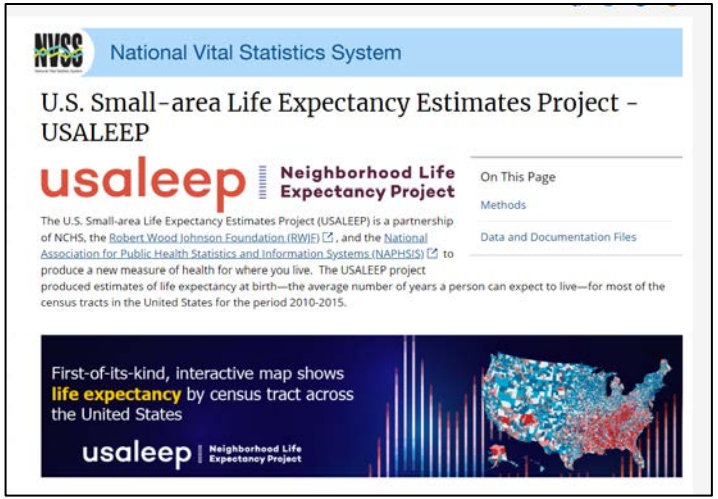
The National Center for Health Statistics (NCHS) received support from the Department of Health and Human Services' Office of the Secretary Patient-Centered Outcomes Research Trust Fund (PCORTF) (1) to link several data sources. The first data source, the National Hospital Care Survey (NHCS), collects ED, inpatient, and outpatient department administrative claims or electronic health records (EHR) data from a sample of participating hospitals (2). These data are not nationally representative. The second source, the National Death Index (NDI), is a centralized database of death records on file in jurisdictional vital records offices for all deaths occurring in the United States, as well as the U.S.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

CDC

# Contextual Information through Geocoding

NCHS Data Linkage Program coordinates geocoding of the Center's survey data and death certificates




Parker, J.D., N. Kravets, and A. Vaidyanathan, Particulate Matter Air Pollution Exposure and Heart Disease Mortality Risks by Race and Ethnicity in the United States: 1997-2009 NHIS with Mortality Followup Through 2011. Circulation, 2017.

# Examples of Research using NCHS Linked Data

*Over **1,000** publications based on NCHS linked data*

## Linked Mortality Data

- Deaths Associated with Underweight, Overweight, and Obesity
- Diet quality and all-cause mortality
- Educational Differentials in US Adult Mortality

## Linked NCHS-CMS Data

- Characteristics of those who chose Medicare Advantage upon Medicare enrollment
- Health service use among the previously uninsured
- Concordance between survey reported childhood asthma and linked Medicaid

## Linked NCHS-HUD Data

- Housing assistance and blood lead levels
- Cigarette smoking and adverse health outcomes among adults receiving federal housing assistance
- Housing assistance associated with insurance rates and unmet medical need

# NCHS is Creating Resources that Support Evidence Building

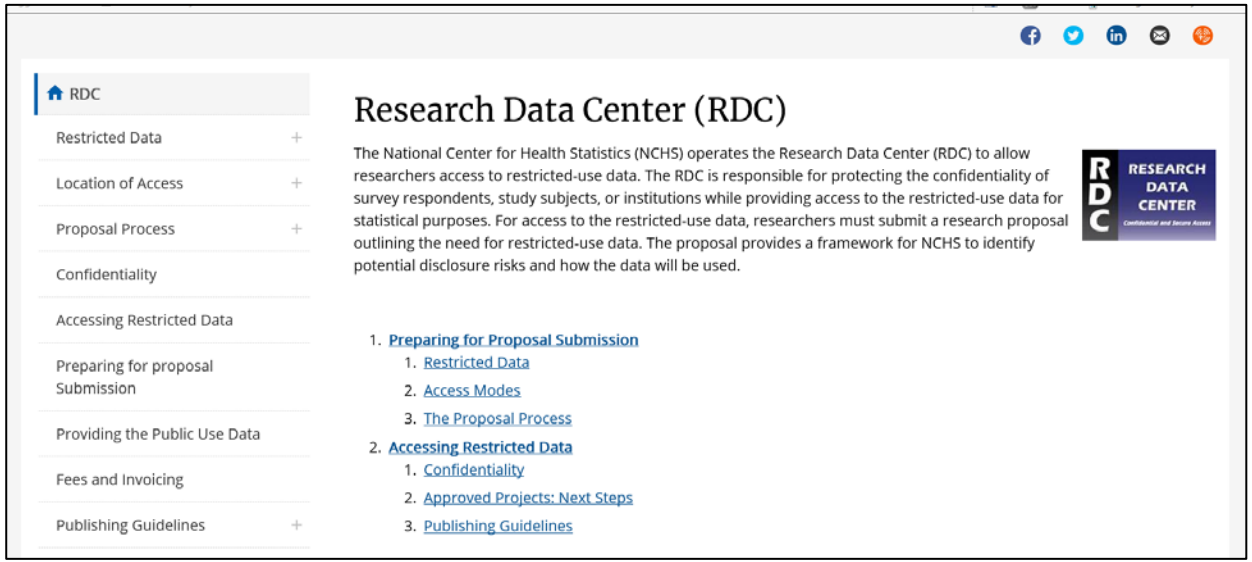
Link NCHS survey data with the National Death Index and other health related administrative data

Provide documentation on linkage methodology, linkage quality and analytic guidelines

Release the curated data files for multiple research projects and to replicate the findings of other researchers

# How to Access NCHS Linked Data: Restricted-use Files

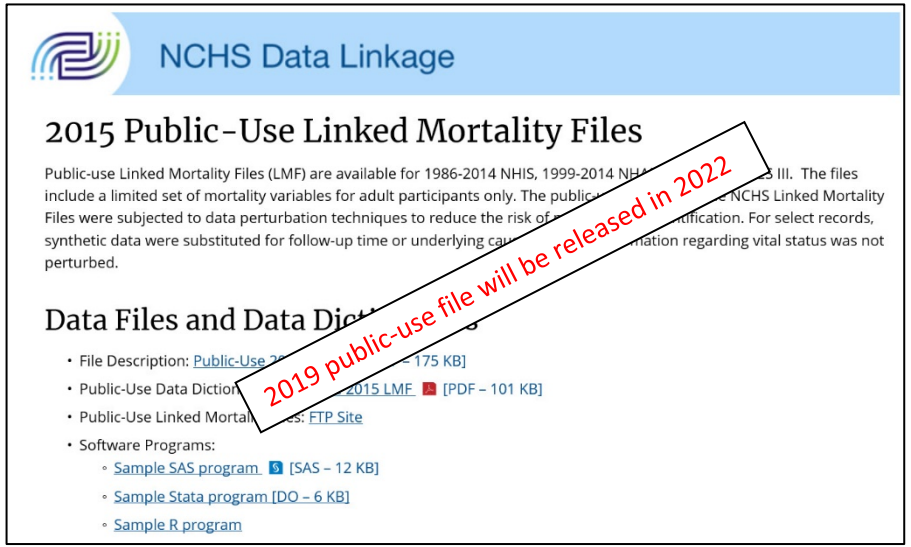
- Most linked data products are available through the NCHS and Federal Statistical Research Data Centers

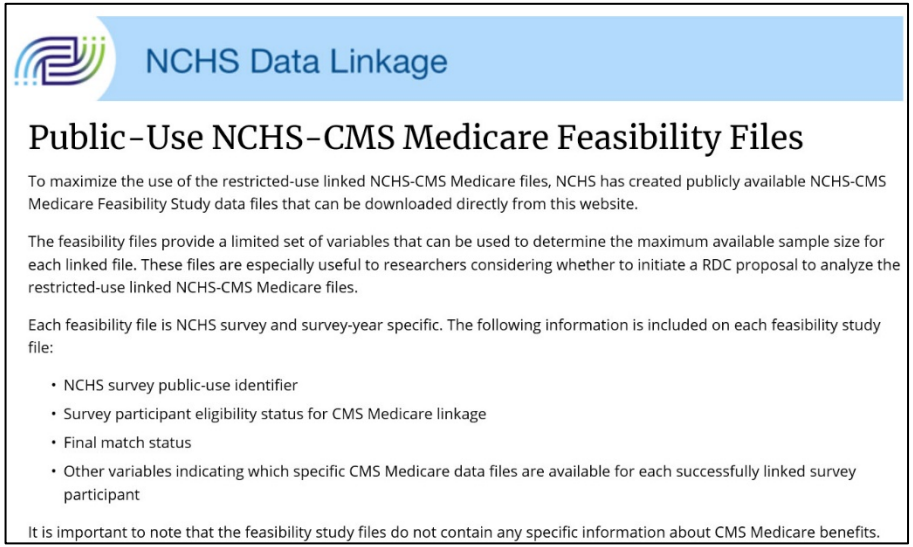



https://www.cdc.gov/rdc/index.htm

# Public-use NCHS Linked Data

- Two types of public-use files released by the NCHS Data Linkage Program:



# Upcoming NCHS Data Linkages

Veterans Affairs

What are the health characteristics, health outcomes, and health care utilization for Veterans within and outside the VA health system?

Anticipated release: early 2022

Medicaid T-MSIS

How do changes in health care policy affect health status for Medicaid recipients?

Anticipated release: early 2022

End Stage Renal Disease

What is the association of dietary intake and patients diagnosed with ESRD?

Anticipated release: September 2021

# NCHS Data Linkage Program Future Directions

- **Machine learning**: Incorporating machine learning techniques to improve linkage accuracy and efficiency

- **Encrypted linkage**: Assessing privacy preserving record linkage algorithms to increase linkage activities across health care spectrum

- **Synthetic data:** Piloting innovative methods for public-use linked data files to improve accessibility and reduce disclosure risk

# More Information

**NCHS Data Linkage Program:**

datalinkage@cdc.gov

www.cdc.gov/nchs/data-linkage

**Lisa Mirel:** LMirel@cdc.gov

**Subscribe to our ListServ** (updates on program including release dates): Send an email message to **list@cdc.gov.** Leave the subject line blank. In the body of the message, type or paste: SUBSCRIBE NCHS-DATA-LINKAGE-PROGRAM lastname, firstname where 'lastname, firstname' is your last and first name.

National Center for Health Statistics
Data Linkage