

NCHS Data Linkage Program

About NCHS

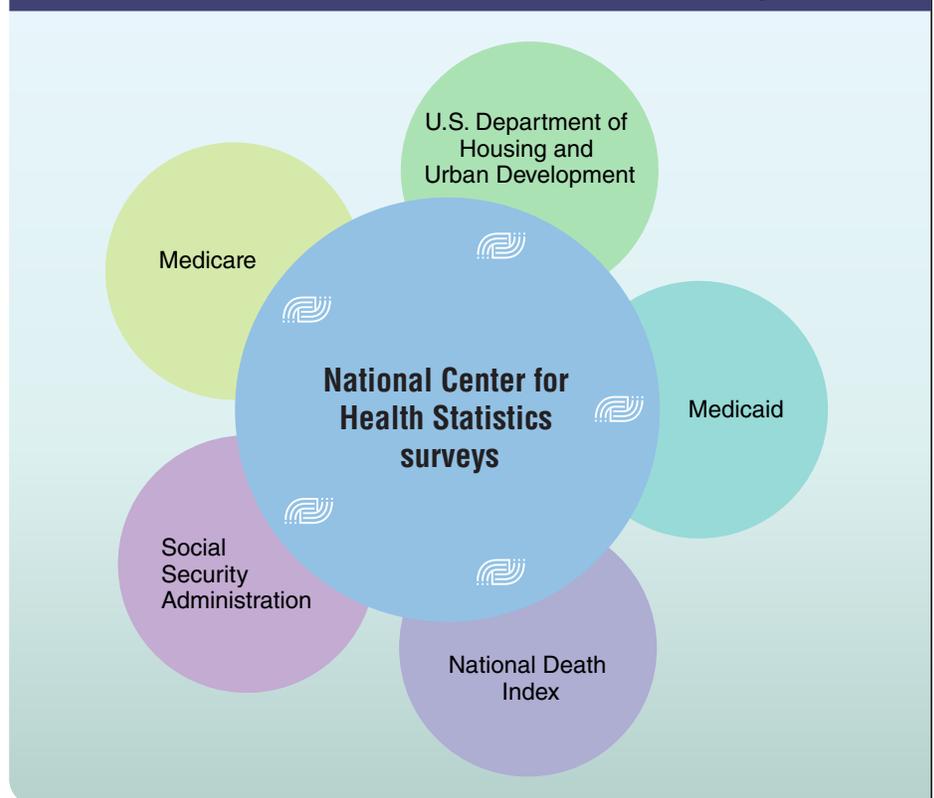
The National Center for Health Statistics (NCHS) is the nation's principal health statistics agency, providing data to identify and address health issues. NCHS compiles statistical information to help guide public health and health policy decisions.

Collaborating with other public and private health partners, NCHS uses a variety of data collection mechanisms to obtain accurate information from multiple sources. This process provides a broad perspective to help understand the population's health, influences on health, and health outcomes.

NCHS data linkage

The NCHS Data Linkage Program is designed to maximize the scientific value of the Center's population-based surveys by linking data from health surveys with data from vital and administrative records. Linked data files enable researchers to examine factors that influence disability, chronic disease, health care utilization, and mortality. Linked data files also maximize scientific value by creating an efficient means to expand the analytic potential of both the survey and administrative data, so enabling analyses that would not be possible with either data source alone.

Current National Center for Health Statistics Data Linkages



NCHS survey data used in linkages include:

- National Health Interview Survey (NHIS)
 - NHIS monitors a broad range of health topics with data collected through personal household interviews.
- National Health and Nutrition Examination Survey (NHANES)
 - NHANES is designed to assess the health and nutritional status of adults and children in the United States. The survey is unique in that it combines interviews with physical examinations.

- National Health Care Surveys
 - The National Health Care Surveys are designed to answer key questions on the use and quality of health care services, including safety and disparities in care provided to population subgroups.

Vital and administrative data sources include:

- National Death Index
 - State vital statistics offices provide records for this centralized database on deaths in the United States.
- Centers for Medicare and Medicaid Services (CMS) Medicare enrollment and claims data, Medicaid Analytic eXtract data
 - CMS provides health insurance and health care coverage for people aged 65 and over and some vulnerable populations in the United States, including low-income children and disabled persons.
- U.S. Department of Housing and Urban Development (HUD)
 - HUD is responsible for overseeing and managing domestic housing programs and policies, including specialized programs for high-needs U.S. populations (e.g., elderly, homeless, and disabled persons).
- Social Security Administration (SSA)
 - SSA administers retirement and disability benefits for elderly and disabled populations in the United States.

Benefits of data linkages

There are many benefits to linking survey and administrative data. The linked data can:

- Provide a longitudinal component to survey data on major diseases, risk factors, and health services utilization through person-level linkages to administrative data on subsequent outcomes
- Reduce research costs by eliminating the need to re-contact survey participants for follow-up information
- Inform survey redesign efforts by supporting research on methodological issues such as the agreement between administrative data and survey questionnaire item responses
- Advance evidenced-based policy making by generating resources that can be used by researchers and policy makers for analytical and statistical purposes

Research examples using linked NCHS data

- Deaths associated with underweight, overweight, and obesity
- Air pollution exposure and heart disease mortality
- Characteristics of beneficiaries who choose Medicare Advantage upon Medicare enrollment at age 65
- Concordance between survey-reported childhood asthma and diagnoses in linked Medicaid administrative data
- Blood lead levels among recipients of federal housing assistance
- Cigarette smoking and adverse health outcomes among adults receiving federal housing assistance

Data access

To protect the confidentiality of survey participants, the linked analytic files are restricted-use and can be accessed only through the NCHS Research Data Centers or Federal Statistical Research Data Centers. Public-use feasibility files, containing a limited set of variables to determine the maximum available sample size for select linked files, are available on the data linkage website.

The linked mortality files consist of public-use and restricted-use files. Public-use linked mortality files contain limited information on cause of death and are available on the data linkage website.

Analyzing linked data

Methods reports are available for all NCHS linked data. They provide information on data sources, linkage methodology, and analytic considerations including descriptions of the variables on the file and tables illustrating the number of survey participants that linked to the vital and administrative data source. NCHS has also created a tutorial for the NHANES-CMS linked data to help researchers navigate analytic approaches when using the linked CMS data (<https://www.cdc.gov/nchs/tutorials/NHANES-CMS/index.htm>).

Future challenges and opportunities

NCHS will continue to develop these resources by exploring opportunities to:

- Expand data availability while continuing to protect confidentiality
- Provide information to inform survey redesign efforts
- Develop resources that can be used to inform policy

