



National Center for Health Statistics

Data Linkage

NCHS Survey Data Linked to the National Death Index (NDI)

Clinton J. Thompson, Ph.D.

NCHS | OAE | SPB

AcademyHealth 2016 Annual Research Meeting
June 26-28
Boston, MA

National Center for Health Statistics

Office of Analysis and Epidemiology



Outline

The Linked Mortality File (LMF) and the National Death Index (NDI)

Creation of the LMF

Content of the LMF

Usage of the LMF



National Center for Health Statistics

Data Linkage

The LMF and the NDI



National Center for Health Statistics

Data Linkage

The LMF and the NDI

NCHS links various surveys with death certificate records from the NDI. This file is known as the LMF.

- ▶ The currently available LMF includes NDI data through 12/31/2011
- ▶ The next LMF expected in 2016 with mortality follow-up through 2015
- ▶ LMFs historically updated every 3-4 years



What is the NDI?



The NDI is a centralized database of U.S. death records gathered from states' vital statistics offices

- ▶ Established to aid health & medical researchers in their research
- ▶ Became operational in 1981
- ▶ Includes death record information for persons dying in the U.S. or a U.S. territory from 1979 onward
- ▶ Death records added annually
- ▶ Detailed information on underlying and multiple cause of death available



National Center for Health Statistics

Data Linkage

Creation of the LMF



National Center for Health Statistics

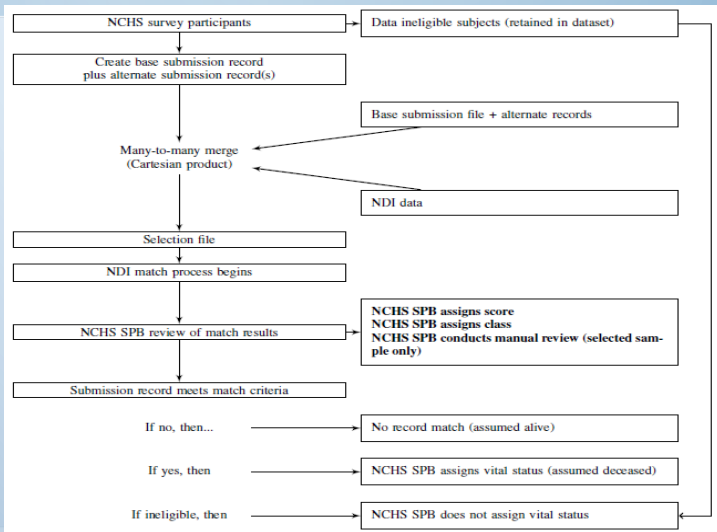
Data Linkage

Which Surveys Link?

- ▶ National Health Interview Survey (NHIS) 1985-2009
(1985-2013)
 - ▶ Supplement on Aging (SOA)
 - ▶ Second Longitudinal Study of Aging (LSOA II)
- ▶ National Health and Nutrition Examination Survey (NHANES) 1999-2010 (1999-2012)
 - ▶ NHANES I Epidemiologic Follow-up Survey (NHEFS)
 - ▶ NHANES III
- ▶ National Nursing Home Survey (NNHS) 1985, 1995, 1997, 2004
- ▶ National Home and Hospice Care Survey (NHHCS) 2007



Linkage Process Overview



National Center for Health Statistics

Data Linkage

Linkage Method

- ▶ Based on framework developed by Fellegi & Sunter (1969) and implemented by NCHS staff working with NDI
- ▶ Score assigned to reflect degree of agreement between NCHS survey submission record and NDI death record
- ▶ Score based on sum of weights for each data item in the record match
- ▶ Weight, w_i , for i^{th} field given by $w_i = \log_2\left(\frac{1}{p_i}\right)$, where p_i is proportion of given characteristic in population



Linkage Eligibility and Matching Criteria

Linkage Eligibility: Requires sufficient data for all elements in one of the three combinations below:

- ▶ SSN, Last Name, First Name
- ▶ SSN, Sex, Month of Birth, Year of Birth
- ▶ Last Name, First Name, Month of Birth, Year of Birth

Matching Criteria: Based on various combinations and subsets of SSN, first/last name (exact & sound-alike), middle initial, month of birth, year of birth (exact & approximate), and father's surname.



Content of the LMF



National Center for Health Statistics

Data Linkage

Restricted-Use File

- ▶ Includes all linked surveys
- ▶ Children and adults included
- ▶ Date of death includes day, month, and year
- ▶ Underlying cause of death for both ICD-9 and ICD-10 codes
- ▶ Data are not perturbed
- ▶ Accessible only in the NCHS Research Data Center (RDC) following submission and approval of research proposal



National Center for Health Statistics

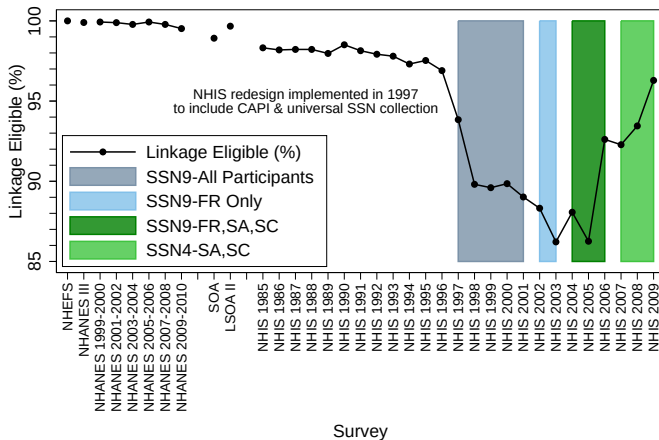
Data Linkage

Public-Use File

- ▶ LSOA II, NHANES III, Continuous NHANES (1999-2010), and NHIS 1986-2009
- ▶ Mortality information for children (<18) suppressed
- ▶ Less detailed death information provided
 - ▶ Date of death limited to quarter and year
 - ▶ Cause of death restricted to leading causes
- ▶ Some data perturbed to reduce risk of participant re-identification



Linkage Eligibility (%)



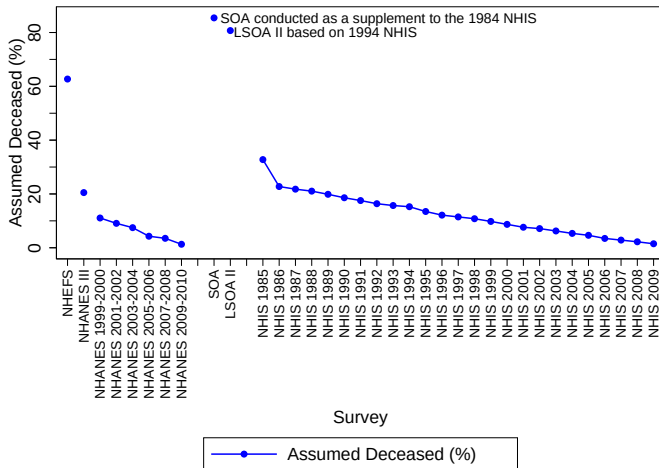
FR-Family Respondent; SA-Sample Adult; SC-Sample Child
CAPI: Computer-Assisted Personal Interviewing



National Center for Health Statistics

Data Linkage

Assumed Deceased (%)



National Center for Health Statistics

Data Linkage

Usage of the LMF



National Center for Health Statistics

Data Linkage

Why Use the LMF?

- ▶ Allows outcomes, e.g. cause of death, to be added to survey data
- ▶ Adds risk-factor and health characteristics to outcomes data
- ▶ Adds a longitudinal component:
 - ▶ Survival
 - ▶ Time to event



Using the LMF

1. Public- or restricted-use?
2. Which NCHS survey(s)?
3. Appropriate merge key/survey identifier (e.g. PUBLICID, SEQN) identified?
4. Correct weight and design variables identified and properly used?
5. Cause of death — desired level of detail?



Items to Consider

- ▶ Previous releases of the LMF should *not* be used. Improvements to matching algorithm and personally identifiable information (PII) make the current (latest) LMF more accurate.
- ▶ There are survey participants in both the 1991 and 1992 NHIS samples so pooling of these two samples should take this into consideration. Refer to the NHIS website for information on how to control for this and other changes in the NHIS survey design.
- ▶ Files can be linked across different administrative sources but this must be requested in a RDC proposal and analyzed in the RDC.



Where Do I Begin?

Visit the NDI mortality data pages of the data linkage website:
<http://www.cdc.gov/nchs/data-linkage/mortality.htm>

National Center for Health Statistics

Data Linkage

What's New

Using Linked Data Products

NDI Mortality Data

Public-Use Files

Restricted-Use Data

Linkage Methods / Analytic Support

 **NCHS Data Linkage**

[CDC](#) > [NCHS](#) > [Data Linkage](#)

NCHS Data Linked to NDI Mortality Files

NCHS has linked various surveys with death certificate records from the National Death Index (NDI). Linkage of the NCHS survey participants with the NDI mortality data provides the opportunity to conduct a vast array of outcome studies designed to investigate the association of a wide variety of health factors with mortality.



National Center for Health Statistics

Data Linkage

Examples of the LMF Used in Research

Pediatrics

March 2013, VOLUME 131 / ISSUE 3

Cardiometabolic Risk Factors Among US Adolescents and Young Adults and Risk of Early Mortality

Sharon Saydah, Kai McKeever Bullard, Giuseppina Imperatore, Linda Geiss, Edward W. Gregg

Osteoporos Int. Author manuscript; available in PMC 2016 Feb 18.

Published in final edited form as:

[Osteoporos Int. 2015 Jan; 26\(1\): 93–102.](#)

Published online 2014 Oct 9. doi: [10.1007/s00188-014-2804-1](#)

PMCID: PMC4757903

NIHMSID: NIHMS758504

Dysmobility syndrome and mortality risk in US men and women age 50 years and older¹

[Anne C. Looker](#), Ph.D



Annals of Epidemiology

Volume 23, Issue 1, January 2013, Pages 31–34



Brief communication

Trends in the coronary heart disease risk profile of middle-aged adults

Ellen Kramarow, PhD^a, James Lubitz, MPH^b, Robert Francis Jr, PhD^a

International Journal of Obesity (2010) **34**, 1044–1050; doi:10.1038/ijo.2010.41; published online 9 March 2010

Weight loss from maximum body weight and mortality: the Third National Health and Nutrition Examination Survey Linked Mortality File

D D Ingram¹ and M E Mussolino²

¹Office of Analysis and Epidemiology, Centers for Disease Control and Prevention, National Center for Health Statistics, Hyattsville, MD, USA

²Division of Cardiovascular Sciences, National Institutes of Health, National Heart, Lung, and Blood Institute, Bethesda, MD, USA

Correspondence: Dr DD Ingram, Office of Analysis and Epidemiology, National Center for Health Statistics, 3311 Toledo Rd., Room 6211, Hyattsville, MD 20782, USA. E-mail: ddingram@cdc.gov



National Center for Health Statistics

Data Linkage

Summary

- ▶ The LMF provides opportunities for new health, policy, and methodology research
- ▶ The public-use LMF is available online but the restricted-use LMF must be accessed through the RDC
- ▶ An updated LMF is expected later this year



National Center for Health Statistics

Data Linkage

Questions? Comments?

Thank you!

Email Contact: `cjthompson1@cdc.gov`

Many thanks to the SPB Linkage Mortality Team:

Jesse Bassich | Jim Brittain | Eileen Call | Adam Fedorowicz |
Cordell Golden | Dedun Ingram | Patsy Lloyd | Lisa Mirel | Keith
Zevallos

