The National Health Interview Survey (1986 - 2004) Linked Mortality Files

Restricted-use File Layout

VARIABLE	DATA TYPE	FIELD DESCRIPTION
PUBLICID	char	NHIS Public-use ID – <u>See Detailed Notes</u>
PUBLICID_2	char	Alternate NHIS Public-use ID for 1997-2003 NHIS years <u>See Detailed Notes</u>
ELIGSTAT	numeric	Eligibility Status for Mortality Follow-up – <u>See Detailed Notes</u>

	1	Eligible
	3	Ineligible

Death Information

MORTSTAT	numeric	Final Mortality Status – <u>See Detailed Notes</u>

0	Assumed alive
1	Assumed deceased
Blank	Ineligible

MORTSRCE_NDI	numeric	Mortality Source: NDI Match – <u>See Detailed Notes</u>

0	No
1	Yes
Blank	Ineligible or assumed alive

MORTSRCE_SSA	numeric	Mortality Source: SSA Information – <u>See Detailed Notes</u>

0	No
1	Yes
Blank	Ineligible or assumed alive

MORTSRCE_CMS	numeric	Mortality Source: CMS Information – <u>See Detailed Notes</u>

0	No
1	Yes
9	Not applicable (1986-1993 NHIS only)
Blank	Ineligible or assumed alive

Restricted-use File Layout

VARIABLE	DATA TYPE	FIELD DESCRIPTION

DODMONTH   numeric   Month of Death

01	January
02	February
03	March
04	April
05	May
06	June
07	July
08	August
09	September
10	October
11	November
12	December
Blank	Ineligible or assumed alive

DODDAY   numeric   Day of Death

01 - 31	Day
Blank	Ineligible or assumed alive

DODYEAR   numeric   Year of Death

1986 - 2006	Year
Blank	Ineligible or assumed alive

AGEDEATH   numeric   Age at Death – [See Detailed Notes]

0 - 113	Age in years
Blank	Ineligible or assumed alive

AGEPRALV   numeric   Age When Last Presumed Alive (as of December 31, 2006)
[See Detailed Notes]

2 - 125	Age in years
Blank	Ineligible or assumed deceased

The National Health Interview Survey (1986 - 2004) Linked Mortality Files

Restricted-use File Layout

VARIABLE	DATA TYPE	FIELD DESCRIPTION

Death Certificate Data

CAUSEAVL   numeric   Cause of Death Data Available – See Detailed Notes

0	No
1	Yes
Blank	Ineligible or assumed alive

ICD_9REV   char   ICD-9 Underlying Cause of Death (1986-1998)
See Underlying and Multiple Causes

001 – 999	Code range
Blank	Cause of death data not available or death after 1998

UCOD_282   char   ICD-9 (1986-1998) Underlying Cause of Death 282 Groups Recode
See Underlying and Multiple Causes

00300 – 35800	Code range
Blank	Cause of death data not available or death after 1998

UCOD_72   char   ICD-9 (1986-1998) Underlying Cause of Death 72 Groups Recode
See Underlying and Multiple Causes

010 – 840	Code range
Blank	Cause of death data not available or death after 1998

UCOD_34   char   ICD-9 (1986-1998) Underlying Cause of Death 34 Groups Recode
See Underlying and Multiple Causes

010 – 370	Code range
Blank	Cause of death data not available or death after 1998

ICD_10REV   char   ICD-10 Underlying Cause of Death (1999-2006)
See Underlying and Multiple Causes

A00 – Y89.9	Code range
Blank	Cause of death data not available or death before 1999

UCOD_358   char   ICD-10 (1999-2006) Underlying Cause of Death 358 Groups Recode
See Underlying and Multiple Causes

001 – 456	Code range
Blank	Cause of death data not available or death before 1999

The National Health Interview Survey (1986 - 2004) Linked Mortality Files

Restricted-use File Layout

VARIABLE	DATA TYPE	FIELD DESCRIPTION

UCOD_39 char ICD-10 (1999-2006) Underlying Cause of Death 39 Groups Recode
See Underlying and Multiple Causes

001 – 042	Code range	
Blank	Cause of death data not available or death before 1999	

UCOD_113 char ICD-10 (all years) Underlying Cause of Death 113 Groups Recode
See Underlying and Multiple Causes

001 – 135	Code range	
Blank	Cause of death data not available	

Multiple Cause of Death Conditions

NUMENTAX numeric Number of Entity Axis Conditions
See Underlying and Multiple Causes

00 – 20	Code range	
Blank	Cause of death code not available	

Variable	Type	Description
COND_1	char	Entity Axis 1st Condition
COND_2	char	Entity Axis 2nd Condition
COND_3	char	Entity Axis 3rd Condition
COND_4	char	Entity Axis 4th Condition
COND_5	char	Entity Axis 5th Condition
COND_6	char	Entity Axis 6th Condition
COND_7	char	Entity Axis 7th Condition
COND_8	char	Entity Axis 8th Condition
COND_9	char	Entity Axis 9th Condition
COND_10	char	Entity Axis 10th Condition
COND_11	char	Entity Axis 11th Condition
COND_12	char	Entity Axis 12th Condition
COND_13	char	Entity Axis 13th Condition
COND_14	char	Entity Axis 14th Condition
COND_15	char	Entity Axis 15th Condition
COND_16	char	Entity Axis 16th Condition
COND_17	char	Entity Axis 17th Condition
COND_18	char	Entity Axis 18th Condition
COND_19	char	Entity Axis 19th Condition
COND_20	char	Entity Axis 20th Condition

VARIABLE	DATA TYPE	FIELD DESCRIPTION
NUMRAXCN	numeric	Number of Record Axis Conditions

See Underlying and Multiple Causes

	00 – 20	Code range
	Blank	Cause of death code not available

RCOND_1	char	Record Axis 1st Condition
RCOND_2	char	Record Axis 2nd Condition
RCOND_3	char	Record Axis 3rd Condition
RCOND_4	char	Record Axis 4th Condition
RCOND_5	char	Record Axis 5th Condition
RCOND_6	char	Record Axis 6th Condition
RCOND_7	char	Record Axis 7th Condition
RCOND_8	char	Record Axis 8th Condition
RCOND_9	char	Record Axis 9th Condition
RCOND_10	char	Record Axis 10th Condition
RCOND_11	char	Record Axis 11th Condition
RCOND_12	char	Record Axis 12th Condition
RCOND_13	char	Record Axis 13th Condition
RCOND_14	char	Record Axis 14th Condition
RCOND_15	char	Record Axis 15th Condition
RCOND_16	char	Record Axis 16th Condition
RCOND_17	char	Record Axis 17th Condition
RCOND_18	char	Record Axis 18th Condition
RCOND_19	char	Record Axis 19th Condition
RCOND_20	char	Record Axis 20th Condition

Survey Data

DOBMONTH	numeric	Month of Birth - See Detailed Notes

	01	January
	02	February
	03	March
	04	April
	05	May
	06	June
	07	July
	08	August
	09	September
	10	October
	11	November
	12	December

VARIABLE	DATA TYPE	FIELD DESCRIPTION
DOBDAY	numeric	Day of Birth - See Detailed Notes
		01 - 31 Day
DOBYEAR	numeric	Year of Birth - See Detailed Notes
		1880 - 2005 Year
INTVMONTH	numeric	Month of Interview

01	January
02	February
03	March
04	April
05	May
06	June
07	July
08	August
09	September
10	October
11	November
12	December

VARIABLE	DATA TYPE	FIELD DESCRIPTION
INTVDAY	numeric	Day of Interview
		01 - 31 Day
INTVYEAR	numeric	Year of Interview
		1986 - 2005 Year
AGEINTV	numeric	Age at Interview - See Detailed Notes
		0 - 113 Age in years

Eligibility Adjusted Weights

WGT_NEW   numeric   Weight Adjusted for Ineligible Respondents - Person-level Sample Weight

  Blank   Ineligible

1986 NHIS: Values for WGT_NEW are equal to the annual final basic weight from the 1986 NHIS public-use file (WTFA).

1987-2004 NHIS: Values for WGT_NEW reflect the eligibility adjusted Person-level Sample Weights created to account for those ineligible for linkage to the NDI.
See Detailed Notes

SA_WGT_NEW   numeric   Weight Adjusted for Ineligible Respondents - Sample Adult Sample Weight

1986-1996 NHIS: The Sample Adult Sample Weight is not applicable for the 1986-1996 NHIS analyses and is therefore left blank. NCHS initiated a redesign of the NHIS questionnaire that was implemented in 1997. The Sample Adult File does not exist prior to the 1997 NHIS.

1997-2004 NHIS: SA_WGT_NEW is available on the file. Blank values represent those ineligible or not a sample adult.
See Detailed Notes

SC_WGT_NEW   numeric   Weight Adjusted for Ineligible Respondents - Sample Child Sample Weight

1986-1996 NHIS: The Sample Child Sample Weight is not applicable for the 1986-1996 NHIS analyses. NCHS initiated a redesign of the NHIS questionnaire that was implemented in 1997. The Sample Child File does not exist prior to the 1997 NHIS.

1997-2004 NHIS: SC_WGT_NEW is available on the file. Blank values represent those ineligible or not a sample child.
See Detailed Notes

PUBLICID – NHIS Public ID Construction

The NHIS Linked Mortality Data Files are person-level files and can be linked to the NHIS public-use files by matching on the unique person-level NHIS Id number. Across the NHIS years included in the linked mortality files, NHIS has changed its construction of a unique person-level Id.

The id number available on the NHIS Linked Mortality Files is consistent with the documentation provided for each NHIS year. Below are instructions on constructing a person-level id from the NHIS public-use files that can be merged with the NHIS Linked Mortality Files. The public id variable on the restricted-use NHIS Linked Mortality Files is in character format with an assigned length of 14.

1986-1994 NHIS
The data items QUARTER x PSU x WEEK x SEGMENT x HOUSEHOLD NUMBER x PERSON NUMBER identify a person within each NHIS year. When combining across NHIS years it is necessary to include the NHIS year as part of the unique person-level id.

Each file is in sequence by the following identifying items that can be concatenated to create the unique person identifier. Note: The NHIS SAS input statements available from the NHIS public-use data website do NOT input all of these variables as character and they must be in character format.

Item	NHIS Public-use File Location	Variable length	Description
Year (2-digit)	3-4	2	Year of interview
Quarter	5	1	Calendar quarter of interview
PSU	6-8	3	Random recode of PSU
Week	9-10	2	Week of interview within quarter
Segment	11-12	2	Segment number
Household number	13-14	2	Household number
Person number	15-16	2	Person number within household

1995-1996 NHIS
The data items HOUSEHOLD ID x PERSON NUMBER identify a person within each NHIS year. When combining across NHIS years it is necessary to include the NHIS year as part of the unique person-level id.

Each file is in sequence by the following identifying items that can be concatenated to create the unique person identifier. Note: the NHIS SAS input statements available from the NHIS public-use data website do NOT input all of these variables as character and they must be in character format.

Item	NHIS Public-use File Location	Variable length	Description
Year (2-digit)	3-4	2	Year of interview
Household ID	5-14	10	Household ID number
Person number	15-16	2	Person number within household

1997-2003 NHIS
The person-level files (Person, Sample Adult, Sample Child, and Immunization) have been sorted by the Person Number (PX) within the Household Serial Number (HHX) and uniquely identify a person within each NHIS year. When combining across NHIS years it is necessary to include the NHIS year as part of the unique person-level id.

Each file is in sequence by the following identifying items that can be concatenated to create the unique person identifier. Note: the NHIS SAS input statements available from the NHIS public-use data website do NOT input all of these variables as character and they must be in character format.

Item	NHIS Public-use File Location	Variable length	Description
Year (4-digit)	3-6	4	Year of interview
Household serial number	7-12	6	Household number
Person number	15-16	2	Person number within household

Note: The family serial number (location 13-14) is not necessary to uniquely identify a person-level record for the years 1997-2003. However, for users who prefer to include the NHIS family serial number in the NHIS public-use Id number, an additional public-use Id, **PUBLICID_2** has been included on the NHIS Linked Mortality Files (only for the NHIS years 1997-2003). Please indicate in your RDC application which public id variable you are providing.

Restricted-use Detailed Notes

To create **PUBLICID_2** concatenate the following identifying items to get the unique person identifier.

Item	NHIS Public-use File Location	Variable length	Description
Year (4-digit)	3-6	4	Year of interview
Household serial number	7-12	6	Household number
Family serial number	13-14	2	Family number
Person number	15-16	2	Person number within household

2004 NHIS

The person-level files (Person, Sample Adult, and Sample Child) have been sorted by combining the Household Serial Number (HHX), Family Sequence Number (FMX) and Person Sequence Number (FPX) and uniquely identify a person within each NHIS year. When combining across NHIS years it is necessary to include the NHIS year as part of the unique person-level id.

Each file is in sequence by the following identifying items that can be concatenated to create the unique person identifier. Note: the NHIS SAS input statements available from the NHIS public-use data website do NOT input all of these variables as character and they must be in character format.

Item	NHIS Public-use File Location	Variable length	Description
Year (4-digit)	3-6	4	Year of interview
Household serial number	7-12	6	Household number
Family sequence number	13-14	2	Family number
Person sequence number	15-16	2	Person number

ELIGSTAT – Eligibility Status for Mortality Follow-Up

NHIS participants are defined as ineligible for mortality linkage if they had insufficient identifying data to create a National Death Index submission record (See Mortality Linkage Eligibility, Tables 1 and 2). Please note that all NHIS participants are included on the linked mortality files.

MORTSTAT – Final Mortality Status

The MORTSTAT variable is NCHS's final determination of vital status and should be used as an outcome variable and to calculate survival. Each NHIS participant who is eligible for mortality follow-up is assigned a vital status code (0 = assumed alive; 1 = assumed deceased).

MORTALITY SOURCE. NCHS has multiple sources of information to determine the final mortality status of a NHIS survey participant. Variables indicating which source or sources were used to determine vital status are available on the linked mortality files and are described below. Any combination of these mortality sources was used to identify NHIS participants as decedents.

MORTSRCE_NDI - Mortality Source: NDI Match

A code 1 indicates that mortality status was ascertained through a probabilistic match to a National Death Index (NDI) record.

MORTSRCE_SSA - Mortality Source: SSA Information

A code 1 indicates that mortality status was obtained from the Social Security Administration (SSA).

MORTSRCE_CMS - Mortality Source: CMS Information

For the NHIS years 1994-2004, a code 1 indicates that mortality status was obtained from the Centers for Medicare and Medicaid Services (CMS). For the NHIS years 1986-1993, CMS is not a source of mortality information. For these years, this variable is coded as not applicable.

AGEDEATH – Age at Death

Misreporting or discrepancies in reported age at interview or date of birth can result in values for age at death that may be inconsistent with interview age. See Analytic Guidelines.

AGEPRALV – Age When Last Presumed Alive (age as of December 31, 2006)

Misreporting or discrepancies in reported age at interview or date of birth can result in values for age when last presumed alive that may be inconsistent with interview age. See Analytic Guidelines.

CAUSEAVL – Cause of Death Data Available

Death certificate data is available only for cases where the variable MORTSTAT=1 and MORTSRCE_NDI = 1. Users should note that it is possible to have an NDI match but no cause of death data available.

SURVEY DATA AGE AND DATE OF BIRTH VARIABLES.
Age at interview (AGEINTV) has been provided on the linked mortality files for the analyst's convenience. Age has not been top-coded.

Complete date of birth information has been provided on the linked mortality files for the analyst's convenience. Year of birth has not been bottom-coded. Where date of birth information is missing on the NHIS public-use files, it has been imputed with month = 06, day = 15, and YEAR = survey year minus the age at interview. In some cases, the date of birth information provided on the NHIS public-use files is inconsistent with date of interview information and there has been no attempt to reconcile this information. Also, imputing

date of birth information for those less than one year of age can create inconsistencies in date of birth information and interview date.

ELIGIBILITY ADJUSTED WEIGHT VARIABLES
For the 1986 NHIS Linked Mortality File, there are no eligibility adjusted sample weights. NCHS recommends using the public-use annual final basic weight (WTFA). For the analyst's convenience, the variables WGT_NEW have been populated with the values of WFTA on the 1986 NHIS Linked Mortality File.

For the 1987-2004 NHIS Linked Mortality Files, NHIS participants classified as eligible for mortality follow-up had their original NHIS sampling weight adjusted to account for those ineligible for linkage to the NDI. Ignoring those ineligible in the NHIS Linked Mortality Files may lead to biased mortality estimates. For 1987-1996 NHIS, there is one eligibility adjusted sample weight, WGT_NEW. For 1997-2004 NHIS, three eligibility adjusted sample weights have been provided: WGT_NEW, SA_WGT_NEW, and SC_WGT_NEW. See Analytic Guidelines.