

Documentation of Mortality Statistics and Cause of Death for MIDUS Core Samples (National and Milwaukee)

This memo accompanies an updated mortality dataset containing 2,459 decedents from the MIDUS Core National and Milwaukee samples that were ascertained via NDI searches up to 2022.

In 2021, all MIDUS Core sample, regardless of prior known decedent status, were submitted to the National Death Index (NDI) to find new decedents, confirm known decedents, and obtain cause of death information. Two NDI Plus¹ searches were conducted: NDI Final Data Search for years 1995 through 2019; and NDI Early Release Search for 2020. Information from 7,699 Core individuals with sufficient minimum matching criteria² were submitted to NDI. Based on the results of the NDI Plus search, 2,124 decedents were ascertained.

In 2023, all MIDUS Core sample were resubmitted to the NDI to obtain updated mortality information. Similar as the NDI Plus Search in 2021, two NDI searches were conducted: NDI Final Data Search for 2020 and 2021; as well as NDI Early Release Search for 2022. Information from 7,699 Core individuals with sufficient minimum matching criteria² were submitted to NDI. The search adds 335 newly identified decedents to the 2,124 decedents in the previous NDI data search in 2021 for a total of 2,459 decedents in the MIDUS Core sample.

The updated Core mortality dataset:

MIDUS_Core_MortalityCauseData_N2459_20230531.sav

Note that the mortality information of deaths through 2021 is considered “complete” due to the NDI Final Data Search, and the mortality information of deaths in 2022 is considered as “close to complete” due to the NDI Early Release Search (95% completion, according to NDI).

Note the changes in cause of death information in the current update. The prior mortality data included only *underlying* cause of death from the NDI Plus. In the current update, *multiple* cause of death are included additionally³. Both underlying cause of death and multiple cause of death are included in two formats: original ICD-9/10 codes from NDI Plus and simplified categories based on the leading cause of death in the U.S. in recent years including heart disease, cancer, chronic lower respiratory disease (CLRD), diabetes, COVID-19, and other disease.

¹ Different from NDI routine searches which only provide potential decedent matches, the NDI Plus search also provides the cause of death (ICD-9/10 Codes) for most of the true match cases.

² To be eligible for an NDI search, each record must contain at least one of the following combinations of data items: (a) First and last name and SSN, (b) First and last name and month and year of birth, (c) SSN and full date of birth and sex.

³ Each death certificate contains a single underlying cause of death and up to 20 additional multiple causes of death. The NDI Plus provides two sets of multiple cause of death fields: entity-axis codes and record-axis codes. The entity-axis codes represent what is actually written in the death certificate by the certifier, but their main limitation is that they do not necessarily reflect the best code for a condition when considered within the context of the medical certification as a whole. Therefore, the Division of Vital Statistics created a new code set called record-axis multiple cause data from the original set of entity codes. The record-axis codes are assigned based on the codes that best describe the overall medical certification portion of the death certificate. They represent a further processing of the entity-axis codes to ensure consistency with the underlying cause data and to make them more amenable to statistical tabulation and analysis. The record-axis codes are what the MIDUS Core Mortality Data provides as multiple cause of death information. For more information about entity-axis and record-axis codes, please refer to the NDI User Guide on pages 58-61 (https://www.cdc.gov/nchs/data/ndi/ndi_users_guide.pdf).

The mortality dataset is a flat file containing 2,459 deceased cases, and it contains 42 variables that can be linked to other Core MIDUS datasets via the M2ID.

The focal variables are:

- DECEASED indicates decedent status.
- DOD_M and DOD_Y indicate the month and year of death.
- UNDERCOD_ICD9 and UNDERCOD_ICD10 contain ICD (International Classification of Diseases) codes of underlying cause of death, which is defined as “(a) the disease or injury which initiated the train of events leading directly to death, or (b) the circumstances of the accident or violence which produced the fatal injury” (World Health Organization; Manual of the International Statistical Classification of Diseases, Injuries, and Causes of Death, based on the recommendations of the Ninth Revision Conference, 1975; Geneva; 1977).
 - Because NDI utilized both ICD-9 (for the deaths occurred between 1979 and 1998) and ICD-10 (for the deaths on 1999 and afterwards) for classification of cause of death, MIDUS mortality data included the two variables, UNDERCOD_ICD9 and UNDERCOD_ICD10:
 - UNDERCOD_ICD9. This variable contains the 9th revision (ICD-9) of the ICD codes and are used to classify the cause of deaths that occurred from 1995 through 1998. Further information on ICD-9 is available from: <https://www.cdc.gov/nchs/icd/icd9.htm>
 - UNDERCOD_ICD10. This variable contains the 10th revision (ICD-10) of the ICD codes and are used by NDI to classify the cause of deaths occurred in 1999 and afterwards. Further information on ICD-10 is also available from: <https://www.cdc.gov/nchs/icd/icd10.htm>
 - For more information about the differences between ICD-9 and ICD-10 codes, see **Table 1** below.
- MULTICOD_1_ICD9 thru MULTICOD_7_ICD9 contain multiple cause of death information for decedents who died prior to 1999. The file does not have MULTICOD_8_ICD9 thru MULTICOD_12_ICD9 because all deaths in those columns occurred in 1999 or afterwards and consequently had only ICD10 codes.
- MULTICOD_1_ICD10 thru MULTICOD_12_ICD10 contain multiple cause of death information for decedents who died in 1999 or afterwards.
- NUM_MULTICOD denotes the total number of multiple cause of death available for the decedent.
- UNDERCOD_C, and MULTICOD_1_C thru MULTICOD_12_C contain categorized version of underlying and multiple cause of death information. The original ICD codes were recoded into six categories based on the leading cause of death in the U.S. in recent years: Heart disease, Cancer, CLRD, Diabetes, COVID-19, and Other disease. The specific ICD-9 and ICD-10 codes that were used to create the categorized version of cause of death variables are shown in **Table 2** below.

Table 1. Key differences between ICD-9-CM and ICD-10-CM and ICD-10-PCS code sets⁴

Differences Between ICD-9-CM and ICD-10 Code Sets		
	ICD-9-CM	ICD-10 code sets
Procedure	3,824 codes	71,924 codes
Diagnosis	14,025 codes	69,823 codes
ICD-10 Code Structure Changes (selected details)		
	Old	New
Diagnosis Structure	ICD-9-CM <ul style="list-style-type: none"> • 3-5 characters • First character is numeric or alpha • Characters 2-5 are numeric 	ICD-10-CM <ul style="list-style-type: none"> • 3-7 characters • Character 1 is alpha • Character 2 is numeric • Characters 3 – 7 can be alpha or numeric
Procedure Structure	ICD-9-CM <ul style="list-style-type: none"> • 3-4 characters • All characters are numeric • All codes have at least 3 characters 	ICD-10-PCS <ul style="list-style-type: none"> • ICD-10-PCS has 7 characters • Each can be either alpha or numeric • Numbers 0-9; letters A-H, J-N, P-Z

Table 2. ICD-9 and ICD-10 Codes Used to Create Categorized Version of Cause of Death Variables

Cause of Death Group	ICD-9 Code	ICD-10 Code
Heart Disease	390-398, 402, 404, 410-424, 425.0-425.4, 425.6-425.9, 426-429	I00-I09, I11, I13, I20-I41, I42.0-I42.5, I42.7-I42.9, I44-I51
Cancer	140-208	C00-C97
CLRD	490-494, 496	J40-J47
Diabetes	250	E10-E14
Covid-19 ^a	Not applicable	U07.1

^a Included since 2020.⁴ Source: https://www.cdc.gov/nchs/icd/icd10cm_pcs_background.htm