National Death Index

User's Guide

U.S. CENTERS FOR DISEASE CONTROL AND PREVENTION

National Center for Health Statistics

Brian C. Moyer, Ph.D., *Director*
Amy M. Branum, Ph.D., *Associate Director for Science*

Division of Vital Statistics

Paul D. Sutton, Ph.D., *Acting Director*
Andrés A. Berruti, Ph.D., M.A., *Associate Director for Science*

NDI DATA USE RESTRICTIONS

The Public Health Service Act (42 U.S.C. 242m) provides in Section 308(d) that the data collected by the National Center for Health Statistics (NCHS), Centers for Disease Control and Prevention (CDC), may be used only for the purpose of health statistical reporting and analysis.

Each National Death Index (NDI) applicant (plus other parties involved in the study or project) provides NCHS with confidentiality assurances that information obtained from NDI will not be published or released in any form if a particular individual or establishment supplying the information or described in it is **identifiable**.

Additionally, the applicant has assured NCHS that the NDI data:

- May only be used for public health and medical studies, PROVIDED that no actions are taken which may directly affect individuals or establishments.
- May not be used for genealogical, legal, administrative, or commercial purposes or to generate profit.
- May not be sent or shared outside of U.S. territorial borders.
- May only be used for the study or project described in the approved NDI application.

As an added precaution, it is strongly recommended that you destroy any identifiable death record information, especially coded causes of death, associated with those NDI record matches that you subsequently determine to be FALSE matches. Such information is of no value to your study activities and is best protected if it is deleted from your computer files or shredded.

It is also important that you notify NCHS whenever your organization receives (or expects to receive) a subpoena or court order for identifiable death record information obtained as a result of using NDI. NCHS will do everything possible to assist in protecting the confidentiality of the identifiable information.

NCHS Assurance Confidentiality and Data Security Provisions

We take your privacy very seriously. All information that relates to or describes identifiable characteristics of a person, a practice, or an establishment will be used only for statistical purposes. NCHS staff, contractors, and agents will not disclose or release responses in identifiable form without the consent of the person or establishment according to section 308(d) of the Public Health Service Act (42 U.S.C. 242m(d)) and the Confidential Information Protection and Statistical Efficiency Act or CIPSEA (44 U.S.C. 3561-3583). In accordance with CIPSEA, every NCHS employee, contractor, and agent has taken an oath and is subject to a jail term of up to 5 years, a fine of up to $250,000, or both if they willfully disclose ANY identifiable information about you. In addition to the above-cited laws, NCHS complies with the Federal Cybersecurity Enhancement Act of 2015 (6 U.S.C. §§ 151 and 151 note), which protects federal information systems from cybersecurity risks by screening their networks.

For the NDI user's data, NDI requires the user to submit the encrypted identifiable data to NDI. To perform the NDI search, an NDI programmer in the NCHS Hyattsville, Maryland, office submits the user's data to the CDC computer center in Atlanta through a Secure Data Network (SDN) and receives the NDI Search Results through the same SDN.

The NDI search is performed on the data network, servers, and software applications that have all been certified (i.e., have a CDC Authorization to Operate or ATO).

Access to the data is restricted to NDI staff. Additionally, once NDI staff determine that an NDI search has been successfully completed, backups of the NDI user's records and of the NDI Search Results are removed from both the server at the CDC computer center in Atlanta and from the NDI programmers' computers in Hyattsville.

How to Use the National Death Index: Steps in the Process

1. Please review the steps in the application process on the NDI website. (Allow approximately 1 to 3 months for your application to be reviewed and approved.)

2. If you have questions while completing a DRAFT of your NDI Application Form, feel free to call NDI staff at 301–458–4444.

3. Your final application will be distributed to the advisory board for review.

4. While your application is being reviewed, you can begin preparing your study subjects' records for submission. Carefully follow the specifications in Chapter 2 of this User's Guide. NOTE: We require that you password-protect your files using any self-decrypting software. Contact NDI staff if you need to make other arrangements for transmitting your data.

5. Assess your results to identify the **true** matches using your own assessment criteria or the suggested NDI assessment criteria. (It is your responsibility to determine which, if any, of the possible NDI matches listed is actually the correct match.) You also have the option of purchasing death certificates on your own directly from state vital statistics offices—for all **true** matches, or just to adjudicate selected questionable matches.

Preparing Your Records: Record Layout and Coding Specifications

1. Records of study subjects must be created in a **standard TEXT FILE** (flat file) format. NOTE: DO NOT put any header lines or other extraneous information in the file. Only DATA should be in each file.

2. Each record must contain **exactly 100 positions** (containing letters, numbers, or blanks) and should have one carriage return (ENTER key) at the end. Limit file sizes to 500,000 records or less; this improves processing speeds. For file sizes larger than 500,000 records, contact NDI staff for current file size requirements.

3. To be **eligible** for an NDI search, each record must contain **at least one** of the following combinations of data items:

   - **FIRST** and **LAST NAME** and **SOCIAL SECURITY NUMBER**
   - **FIRST** and **LAST NAME** and **MONTH** and **YEAR OF BIRTH**
   - **SOCIAL SECURITY NUMBER** and full **DATE OF BIRTH** and **SEX**

   These combinations of items are part of the NDI Edit requirements that determine whether a user's record has sufficient data to qualify for a search. These combinations should not be confused with the NDI matching criteria described on page 1 of Chapter 4.

4. **DO NOT USE ANY DELIMITERS SUCH AS TABS, COMMAS, OR ANY OTHER PUNCTUATION TO SEPARATE THE DATA ELEMENTS.**

5. Each data element must be entered in its specified positions. See Exhibit 1 for sample user records in the required record layout.

6. Refer to Exhibit 2 for the file format and coding specifications for each data element.

7. Coding instructions for **NAMES** appear in Exhibit 2A.

8. State codes appear in Exhibit 2B.

9. A separate **NDI Transmittal Form** must accompany **each** file submitted for an NDI search. A blank NDI Transmittal Form(s) containing your assigned NDI search numbers will be sent to you with the letter approving your NDI search. NOTE: A SEPARATE FILE must be created for each type of search: routine, unknown, known, and certificate, and similarly for the Early Release File.

10. Submit your records to us. Use your assigned NDI numbers in your file names. **NCHS requires that you password-protect your file(s) using any self-decrypting software.**

11. Send us your files, NDI Transmittal Form(s), fee worksheet, and proof of payment. Refer to "How to Use the National Death Index: Steps in the Process." Proof of payment must be received by NCHS before you receive the NDI Search Results.

12. E-mail ndi@cdc.gov to confirm that your file(s) was received. Include your NDI number and the date the file was sent. Be sure to include the password(s) to be used for opening your password-protected file(s).

NDI User Records, Sample Format

Create your records in a standard **text file** format with **100 positions** per record, with no tabs, commas, delimiters, or other punctuation. Statistical software such as SAS is recommended to create a text file and to position each data element in the exact record positions described below. DO NOT include any header lines or other extraneous information in the file. Only DATA should be in each file. If you are submitting only a small number of records and plan to key your data into a text file, text file software such as **Microsoft Notepad** is recommended. In Notepad you can click on VIEW and activate the STATUS BAR, which enables you to see the line and column numbers at the bottom of the screen. Change the font to a proportionately spaced font such as Lucida Console, which will ensure that each data item aligns properly. Make sure you create a separate file for each type of search: routine, unknown, known, and certificate, and similarly for the Early Release File.

Positions of NDI data items

(1-20)	LAST NAME
(21-35)	FIRST NAME
(36)	MIDDLE INITIAL
(37-45)	SOCIAL SECURITY NUMBER
(46-47)	MONTH OF BIRTH
(48-49)	DAY OF BIRTH
(50-53)	YEAR OF BIRTH
(54-71)	FATHER'S SURNAME
(72)	AGE UNIT (at death)
(73-74)	NUMBER OF AGE UNITS (at death)
(75)	SEX
(76)	RACE
(77)	MARITAL STATUS
(78-79)	STATE OF RESIDENCE
(80-81)	STATE OF BIRTH
(82-91)	CONTROL/ID NUMBER
(92-97)	OPTIONAL USER DATA
(98-100)	BLANK FIELD

Sample format for 100-position user records (data are hypothetical)

```
         1         2         3         4         5         6         7         8         9         0
12345678901234567890123456789012345678901234567890123456789012345678901234567890123456789012345678901234567890
WASHINGTON          GEORGE         T28516085604081731WASHINGTON        075121013059A536a58 060183
WASHINGTON          MARTHA         Q34465413601211922Adams             075121013059387     052880
Franklin            Benjamin       D28878210405221701Franklin          10111216NH49596DUP  120998
ROOSEVELT           ELEANOR         2627021711123192 6ARIAS            029211NVCA62345     122296
Barton              Clara          A34445464407041961Brown             035222MDMD59387     100889
```

NOTE: Please limit file sizes to 500,000 records or less; this improves processing speeds. For file sizes greater than 500,000 records, please contact NDI staff for current file size requirements.

NDI User File Format and Coding Specification

DATA ELEMENTS	POSITIONS	CODE STRUCTURE		CODING INSTRUCTIONS
1. Name of person in study group				
Last Name	1–20	Alpha, left justified		* See EXHIBIT 2A for instructions on coding surnames, first names, and middle initials. * Both LAST name and FIRST name must be provided or the record will automatically be rejected, unless the record contains a Social Security Number, a date of birth, and a sex code. An initial in the FIRST name field is permitted.
First Name	21–35	Alpha, left justified		
Middle Initial	36	Alpha or blank		
2. Social Security Number				
	37–45	Alpha/Numeric or blank		* Must have 9 digits. Do not drop any leading zeroes. * If less than 9 digits or a partial number, leave the field blank. * If more than 9 digits, enter the first 9 digits. * If two SOCIAL SECURITY NUMBERS, enter the first or create a duplicate record. * If the number is unknown or not provided, leave the field blank. * If SOCIAL SECURITY NUMBER is not provided, the MONTH and YEAR OF BIRTH must be provided or the record will automatically be rejected.
3. Date of birth				
Month	46–47	January	01	* The months of January–September and the days 1–9 must have leading zeros. * If either MONTH or YEAR of birth is not provided, the SOCIAL SECURITY NUMBER must be provided or the record will automatically be rejected.
		February	02	
		March	03	
		April	04	
		May	05	
		June	06	
		July	07	
		August	08	
		September	09	
		October	10	
		November	11	
		December	12	
Day	48–49	Days	01–31	
		Unknown or not stated	Blank or 99	
Year	50–53	1850–20xx Blank or 9999 if unknown		

NDI User File Format and Coding Specification—Con.

DATA ELEMENTS	POSITIONS	CODE STRUCTURE		CODING INSTRUCTIONS
4. Father's surname				
	54–71	Alpha, left justified or blank if unknown		* Follow the coding instructions for surnames as specified in EXHIBIT 2A. * For females, provision of the FATHER'S SURNAME is encouraged. * For males, the FATHER'S SURNAME is preferred only when it differs from the LAST NAME.
5. Age at death				
Unit	72	Year: Less than 100	0	* Code the ACTUAL age at death if it is already known. An estimated age at death may be used only when deaths for one year are being searched. * The "unit" field identifies the type of units for which age at death is measured. Example: Age 65 Unit = 0 Number of units = 65 Example: Age 105 Unit = 1 Number of units = 05 Example: Age 8 months Months unit = 2 Number of units = 08 Example: Age unknown Unit = blank or 9 Number of units = blank or 99
		Years: 100 or more	1	
		Months	2	
		Weeks	3	
		Days	4	
		Hours	5	
		Minutes	6	
		Age unknown	Blank or 9	
Number of units	73–74	Enter age	00–99	
		Age unknown	Blank or 99	
6. Sex				
	75	Male	1 or M	
		Female	2 or F	
		Unknown	Blank or 9	

NDI User File Format and Coding Specification—Con.

DATA ELEMENTS	POSITIONS	CODE STRUCTURE		CODING INSTRUCTIONS
7. Race				
	76	White	1*	* White (includes Mexican, Puerto Rican, and all Caucasian)
		Black	2	
		Indian	3*	* Indian (includes American, Alaska Native, Canadian, or Mexican Indian; Eskimo; and Aleut)
		Chinese	4	
		Japanese	5	
		Hawaiian	6*	* Hawaiian (includes part-Hawaiian)
		Other non-White	7*	* Other non-White (includes Cajun and Creole)
		Filipino	8	
		Other Asian or Pacific Islander	0	
		Unknown, not stated, or not classifiable	Blank or 9	
8. Marital status				
	77	Never married/ single	1*	* Other Entries Annulled 1 Separated 2 Common-law marriage 2
		Married	2*	
		Widowed	3	
		Divorced	4	
		Unknown	Blank or 9	
9. State of residence				
	78–79	ALPHA and/or NUMERIC codes		* See EXHIBIT 2B for the ALPHA and/or NUMERIC codes that may be used.
		States	01–51	* Insert last known STATE of residence. If provided by the NDI user, it will be matched against the state of residence as reported on the death certificate.
		Puerto Rico	52	
		U.S. Virgin Islands	53	
		Guam	54	
		Canada	55	
		Cuba	56	
		Mexico	57	
		Remainder of world	59	
		Unknown	Blank or 99	

NDI User File Format and Coding Specification—Con.

DATA ELEMENTS	POSITIONS	CODE STRUCTURE	CODING INSTRUCTIONS
10. State of birth			
	80–81	(Same code structure as "State of residence" above)	* See EXHIBIT 2B for the ALPHA and/or NUMERIC codes that may be used.
11. Identification (ID), control, or sequence number (OPTIONAL)			
	82–91	Any combination of ALPHA/NUMBERIC, left justified. If no numbers are assigned, leave this field BLANK.	* This is an OPTIONAL field to assist NDI users in identifying the records they submit. Insert identification numbers or control numbers assigned to people in the study OR assign "sequence numbers" based on the order that records appear in the file to be submitted to NCHS. * If duplicate or alias records are submitted, a flag could be put at the end of the control number (e.g., D1, D2, DUP1, or DUP2). * NOTE: The computer output will include the control or sequence number for each user record. NCHS will assign a sequence number whenever this field is left blank (e.g., ID00000001, ID00000002, ID00000003).
12. Optional user data			
	92–97	Any combination of ALPHA/NUMBERIC, left justified If this field is not used, leave it BLANK	This OPTIONAL field may be used for any additional information on study subjects. Examples of possible uses are: * Date or year of last contact (e.g., 011584, or 84, or 1984). * For KNOWN decedents, date or year of death (e.g., 023091, or 91, or 1991). * To flag TEST RECORDS of known decedents (e.g., date of death, or D, or DEAD) or deaths prior to 1979 (e.g., PRE 79). * Coded information to identify special subgroups of study subjects. * To flag DUPLICATE or ALIAS records. The NDI computer output will contain whatever the user decides to include in this field
13. Blank field			
	98–100		* The user can include data in this field.

Coding Instructions for Names

Surnames

1. If the entire surname will not fit in the 20-position field, truncate it.

2. If the surname has a space or apostrophe following a prefix, such as Van Braun, Mac Pherson, O'Toole, or O'Mally, the NDI edit program will remove the space and the apostrophe between the prefix and the name.

3. If more than one surname is given separated by a space or hyphen, type each name in the surname field but remove the space or the hyphen. The NDI edit program will also remove such spaces or hyphens on the user's record and put the two names together. You have the option of creating alias records for each of the last names. **Note that only the first five alias records will be free.**

4. If a subject's full name appears to contain a MAIDEN name (for example, Jane Karen SMITH-Lee or Jane SMITH Lee), it is important for the maiden name also to be entered separately in the FATHER'S SURNAME field to enhance the possibility of a match. Listed below are suggestions of how user records may be prepared to cover the two different situations:

   Example: Jane Karen SMITH-Lee

   Key First name: JANE
   Middle initial: K
   Last name: SMITHLEE
   Father's surname: SMITH

   Example: Jane SMITH Lee

   Key First name: JANE
   Middle initial: S
   Last name: LEE
   Father's surname: SMITH

5. If more than one surname is given and you are uncertain as to which surname to put in the last name field, consider the merits of creating a duplicate or alias record(s), *one record for each surname (see item 3)*. This may enhance the effectiveness of the NDI file search, especially for records without Social Security numbers. You may want to flag such duplicate records by inserting codes of your own choosing in either the control number field (positions 82–91) or the optional user data field (positions 92–97).

6. Beginning with 2010 deaths, whenever a state's NDI death record has two last names separated by a dash (-), the NDI software will remove the dash on the NDI record and merge the two names. The NDI software will then create two alias records, one with each last name. In such instances, three NDI records will have the same death certificate number (and the letter "A" should appear to the left of the death certificate number if it is the alias record that matched the user's record).

7. Suffixes or generational identifiers such as II, III, Jr., or Sr. will be removed by the NDI edit program.

8. For ***religious*** names such as Sister Mary Lawrence, enter LAWRENCE as the last name, SISTER MARY as the first name, and a BLANK for middle initial.

Coding Instructions for Names—Con.

First names and middle initials

1. If the entire first name will not fit in the field, truncate it.

2. If two first names and a middle name are given, type both first given names (with or without a space between each name) in the first name field and type the middle initial in the middle initial field (truncate).

   Example: MARY ANN LOUISE Jones

   Key First name: MARY ANN
   Middle initial: L

   Example: MARY ANN Jones (Treat ANN as the middle name unless your records indicate that ANN is part of the first name.)

   Key First name: MARY
   Middle initial: A

3. If a person (a) goes by two distinctly different first names, (b) has a first name and a nickname, or (c) uses the shortened form of a presumed legal name (for example, Bob for Robert, Liz for Elizabeth, or Ben for Benjamin), consider the merits of creating duplicate or alias record(s), *one record for each first name*. This may enhance the matching effectiveness of the NDI file search, especially for records without Social Security numbers. You may want to flag such duplicate records by inserting codes of your own choosing in either the control number field (positions 82–91) or the optional user data field (positions 92–97).

4. If entries have a first initial and a middle name, enter an initial for the first name and an initial for middle name.

5. If entries have multiple middle names or initials, enter only the first initial (for example, for Robert M.L. Jones, type "M" as the middle initial and drop the "L").

6. For infant deaths, names such as "Baby Girl" Jones should be entered as Baby Girl for first name and Jones for last name.

State (or Foreign Country) of Residence, Birth, and Death

State[1] or country	Alpha code	Numeric code	State[1] or country	Alpha code	Numeric code
Alabama	AL	01			
Alaska	AK	02	New Mexico	NM	32
Arizona	AZ	03	New York	NY	33
Arkansas	AR	04	North Carolina	NC	34
California	CA	05	North Dakota	ND	35
Colorado	CO	06	Ohio	OH	36
Connecticut	CT	07	Oklahoma	OK	37
Delaware	DE	08	Oregon	OR	38
District of Columbia	DC	09	Pennsylvania	PA	39
Florida	FL	10	Rhode Island	RI	40
Georgia	GA	11	South Carolina	SC	41
Hawaii	HI	12	South Dakota	SD	42
Idaho	ID	13	Tennessee	TN	43
Illinois	IL	14	Texas	TX	44
Indiana	IN	15	Utah	UT	45
Iowa	IA	16	Vermont	VT	46
Kansas	KS	17	Virginia	VA	47
Kentucky	KY	18	Washington	WA	48
Louisiana	LA	19	West Virginia	WV	49
Maine	ME	20	Wisconsin	WI	50
Maryland	MD	21	Wyoming	WY	51
Massachusetts	MA	22	Puerto Rico	PR	52
Michigan	MI	23	U.S. Virgin Islands	VI	53
Minnesota	MN	24	Guam	GU	54
Mississippi	MS	25	Canada	CN	55
Missouri	MO	26	Cuba	CU	56
Montana	MT	27	Mexico	MX	57
Nebraska	NE	28	Remainder of world	RW	59
Nevada	NV	29	American Samoa	AS	60
New Hampshire	NH	30	Northern Marianas Islands	MP	69
New Jersey	NJ	31	Unknown		99

[1]Includes District of Columbia, Puerto Rico, U.S. Virgin Islands, and Guam.

Your NDI Results

This chapter describes the different files you will receive in the NDI Search Results. Before you begin working with these files, NCHS strongly recommends that you also review Chapter 4 for suggestions on how to assess your results.

Overview of NDI Output Files

- The NDI system generates **eight basic output files** for each NDI search. If you opt for NDI Plus service, you will receive **two additional files** containing the cause-of-death codes for your closer or more likely NDI matches. These two files also contain complete lists of the ICD–9 and ICD–10 cause-of-death codes.
- Five of the 10 files are in PDF table formats and five are in standard text formats. All are stored and labeled "NDI Search Results." A brief description of each NDI output file is presented in Exhibit 3.
- The following five NDI output files are in PDF format to be **printed** at your discretion:
  - ◊ User File Edit Results [**EDITS**]
  - ◊ Summary Retrieval Statistics [**SUMMARY**]
  - ◊ NDI Retrieval Report [**REPORT**]
  - ◊ Cause-of-Death Report [**PRTCAUSE**] (up to 500 records; a maximum of 50 records can be printed)
- After you review the content and size of each of these files, decide which you want to print, either partially or completely. When printing these files, use the **landscape** format and, if necessary, reduce the size of your margins or fonts to capture all information.
- The following two files are in text file format for those users who submit a large number of records and who prefer to use their own software to assess the NDI results:
  - ◊ Combined File of Matching User and NDI Records [**COMBINED**]
  - ◊ Cause-of-Death File [**CAUSE**]
- The files labeled **MATCH**, **NOMATCH**, and **REJECTS** contain only the 100-character records that you submitted. These three files are in text file format. Your records are distributed among these files depending on whether they were involved in at least one NDI match, not involved in any matches, or rejected by the NDI edit program.
- Your results will also contain the file you submitted for the NDI search. Your identifying information and the NDI Search Results are backed up on CDC's data file for 60 days and then erased.
- The remainder of this chapter describes each of the 10 NDI output files.

NDI Output File Descriptions

Each file name begins with a user's NDI Search Number, followed by a period and a file name identifier (e.g., 2012-X099.EDITS, 2012-X099.SUMMARY). The last column of the table below indicates which files are **PDF** files and which are **text** files. NOTE: Each file in the NDI Search Results may not always have the PDF or text file extensions appearing at the end of the file. **If you submit a large volume of records, the following three files may contain only a partial listing of the search results: EDITS, REPORT, and PRTCAUSE.**

FILE NAME	FILE DESCRIPTION	FILE TYPE
EDITS	**Edit Results [EDITS]**—List of rejected user records and statistics on the completeness and quality of each data item.	PDF
SUMMARY	**Summary Retrieval Statistics [SUMMARY]**—Statistics on the number of matches, multiple matches, and matches by state of death.	PDF
REPORT	**NDI Retrieval Report [REPORT]**—User records and corresponding NDI record matches, in preprinted format with column headings. Intended solely for the visual inspection of matches.	PDF
COMBINED	**Combined File of Matching User and NDI Records [COMBINED]**—User record information appears in positions 1–100; NDI record information (from the Retrieval Report) appears in positions 101–178. A separate combined record is created for each multiple match. This file is intended solely for designing computer programs to assess a large number of matches. It contains probabilistic scoring in positions 172–178.	Text: 178 positions
CAUSE *	**Cause-of-Death File[CAUSE]**—Contains only those possible NDI record matches that are ranked first or assigned a high probabilistic score. Positions 1–178 are identical to those in the COMBINED file. Positions 180–438 contain the coded causes of death.	Text: 438 positions
PRTCAUSE *	**Cause-of-Death Report [PRTCAUSE]**—Intended to be *printed* solely for the visual inspection of the coded causes of death for the same NDI record matches included in the CAUSE file (up to 500 records; a maximum of 50 records can be printed).	PDF
MATCH	**MATCHING User Records (MATCH)**—User records involved in matches, in the same record format as submitted by the user.	Text: 100 positions
NOMATCH	**NONMATCHING User Records (NOMATCH)**—User records not involved in matches with any NDI records, in the same record format as submitted by the user.	Text: 100 positions
REJECTS	**REJECTED User Records (REJECTS)**—User records rejected by the NDI edit program and not included in the NDI search, in the same record format as submitted by the user.	Text: 100 positions

***PROVIDED ONLY TO NDI USERS WHO REQUEST THE NDI PLUS SERVICE TO OBTAIN CAUSE-OF-DEATH CODES.**

User File Edit Results [EDITS]

All names and other data elements used in this exhibit are hypothetical.

NDI **EDIT RESULTS** CDC

NDI User :	University of Southern North Dakota
Address :	12345 South Dakota Way
	Southern, ND 54321
Phone No :	999-888-7777
E-mail :	Efg@snd.edu

NDI Search Number :	2001-X516

EDIT STATISTICS

	"TOTAL" RECORDS EDITED				RECORDS REJECTED			
	VALID DATA		INVALID DATA	MISSING DATA	VALID DATA		INVALID DATA	MISSING DATA
Data Items	Number	Percent*	Out Of Range	Blank/ Unknown	Number	Percent**	Out of Range	Blank/ Unknown
Last Name	41	78.85	0	11	5	45.45	0	6
First Name	43	82.69	0	9	6	54.54	0	5
Middle Initial	21	40.38	0	31	3	27.27	0	8
Social Security Number	34	65.38	5	13	7	63.64	0	4
Birth Month	47	90.38	0	5	8	72.73	0	3
Birth Day	41	78.85	0	11	6	54.55	0	5
Birth Year	39	75.00	8	5	7	63.64	0	4
Father's Surname	5	9.62	0	47	2	18.18	0	9
Age Unit	37	71.15	0	15	7	63.64	0	4
Age - Number of Units	23	44.23	14	15	7	63.64	0	4
Sex	34	65.38	0	18	6	54.54	0	5
Race	32	61.54	0	20	8	72.72	0	3
Marital Status	0	0.00	0	52	8	0.00	0	11
State of Residence	35	67.31	0	17	2	72.72	0	3
State of Birth	0	0.00	0	52	0	18.18	0	9

*DENOMINATOR :	Total Records Edited	=	52
**DENOMINATOR :	Records Rejected & Listed	=	11
	Records Passed To NDI Retrieval	=	41

User File Edit Results [EDITS]—Con.

All names and other data elements used in this exhibit are hypothetical.

LIST OF USER RECORDS REJECTED BY NDI EDIT PROGRAM

(List is Limited to a Maximum Of 250 Records)

The following records do not contain minimum data needed for a search:

Control or ID Number	Last Name	First Name	Middle Initial	Social Security Number	Birth Date	Father's Surname	Age	Sex	Race	Marital Status	Res. State	Birth State
12834	WILSON	-	R	566666666	-	-	35	-	1	-	33	-
13278	-	ROBERT	-	623456789	-	-	45	F	1	-	33	45
15893	SMITH	-	-	345678901	01/23/1909	-	75	-	2	-	33	-
16444	-	-	-	987654321	01/ /1909	-	90	F	1	-	33	33
16947	-	-	-	123456789	/23/1909	-	55	F	1	-	33	-
18456	WILLIAMS	-	-	234567890	01/23/1909	-	58	-	2	-	33	-
20654	-	MARIA	F	79588607	01/23/1909	-	44	F	1	-	33	-
22345	ANDERSON	JANE	-	-	09/02/	-	-	F	1	-	-	-
27222	JONES	V	H	-	09/02/	-	-	-	-	-	-	-
34567	-	MARY	-	-	09/ /1899	FRANKLIN	-	-	-	-	-	-
35678	-	JOHN	-	-	09/ /1945	DAVIS	-	M	-	-	22	-

Edit Results [EDITS]

1. The **EDITS** file is intended only for you to view or print. It is a PDF file. Use the landscape print format and reduce margin widths if necessary. A sample of the edit results is shown in Exhibit 4.

2. It is important to review your Edit Results first, to confirm that each of the data items you submitted met NDI specifications and to determine if any of your records were rejected, and why. The edit results will provide you with the following information:

   - A listing of your rejected records with a reason for the rejection.
   - The total number and percentage of rejected records.
   - Statistics on the overall completeness of each data item.
   - Statistics on the accuracy of each data item (i.e., how many records contained data that were out of range).

3. The file(s) of records you submit for an NDI file search are first run against the NDI edit program to determine whether your file conforms to NDI file format requirements and coding specifications, described in Chapter 2. If your data file has significant problems, the data will be returned to you for revision and resubmission at no extra charge (up to 5 reruns are free).

4. **User records rejected by the edit program are not eligible for an NDI search.** To be eligible for an NDI search, a record must contain **at least one** of the following combinations of data items:

   - **FIRST** and **LAST NAME** and **SOCIAL SECURITY NUMBER**
   - **FIRST** and **LAST NAME** and **MONTH** and **YEAR OF BIRTH**
   - **SOCIAL SECURITY NUMBER** and **DATE OF BIRTH** and **SEX**

Summary Retrieval Statistics [SUMMARY]

The **SUMMARY** file is a PDF file intended only for you to view or print. Use the **landscape** print format and reduce margin widths if necessary.

This file contains a two-page report (see Exhibit 5) that provides the following summary information about your NDI file search:

1. Your assigned NDI search number and the date of the search. (The search date is not shown in Exhibit 5.)

2. The year(s) of death that your records were searched against and the total number of **NDI records** searched.

3. The number of NDI RECORDS involved in **possible** record matches (i.e., satisfying any one of the seven NDI matching criteria). Also shown is the number of possible record matches that were EXACT matches (i.e., ALL items provided on the user record agreed exactly with the items on the matching NDI record).

4. The total number of **user records** you submitted, the number of your records involved in possible matches, the number of nonmatching records, and the number of your records rejected by the NDI edit program.

5. The extent to which user records were involved in **multiple** matches with **NDI records** (i.e., how many user records matched with only one NDI record, how many matched with exactly two NDI records, and so on, up to how many user records matched with 11 or more NDI records). **NOTE**: Data on the number of multiple possible matches are intended only to show you how many possible matches can be generated by the NDI's liberal matching criteria. The matching criteria are designed to maximize the number of TRUE matches that are generated (listed), but this process requires that you expend time and effort to carefully assess your NDI results to ensure that you select only the correct matches.

6. The number of **NDI records** involved in possible matches by state of death.

Exhibit 5

Summary Retrieval Statistics [SUMMARY]

All data elements used in this exhibit are hypothetical.

NDI **SUMMARY RETRIEVAL STATISTICS** CDC

Year(s) searched : 1979-2007

NDI search number: 2003-X456

Records on the National Death Index file:

Total NDI records searched :	66,645,785
No. of NDI records involved in possible record matches :	385
All items matched :	51
Other possible matches :	334

Request records submitted by the user:

Total user request records submitted :	180
No. of user records involved in matches:	159
No. of non-matching user records :	21
No. of rejected user records :	0

NUMBER OF USER REQUEST RECORDS INVOLVED IN MATCHES WITH ONE OR MORE NDI RECORDS

Possible NDI record matches	Number of user request records
1 NDI Records	100
2 NDI Records	22
3 NDI Records	15
4 NDI Records	8
5 NDI Records	0
6 NDI Records	1
7 NDI Records	3
8 NDI Records	2
9 NDI Records	1
10 NDI Records	0
11+ NDI Records	7
Total user records involved in matches	**159**

Summary Retrieval Statistics [SUMMARY]—Con.

All data elements used in this exhibit are hypothetical.

NDI **SUMMARY RETRIEVAL STATISTICS** CDC

NUMBER OF NDI RECORDS INVOLVED IN POSSIBLE RECORD MATCHES WITH THE USER'S REQUEST RECORDS BY REGISTRATION AREA OF DEATH

States	Total Possible Record Matches	Matches On All Items Provided*	Other Possible Record Matches**
Alabama	7	1	6
Alaska	0	0	0
Arizona	7	1	6
Arkansas	6	2	4
California	35	2	33
Colorado	4	1	3
Connecticut	0	0	0
Delaware	2	0	2
District of Columbia	3	0	3
Florida	18	1	17
Georgia	7	0	7
Hawaii	1	0	1
Idaho	2	0	2
Illinois	23	5	18
Indiana	11	2	9
Iowa	2	1	1
Kansas	1	0	1
Kentucky	8	1	7
Louisiana	11	1	10
Maine	2	0	2
Maryland	5	0	5
Massachusetts	6	1	5
Michigan	10	0	10
Minnesota	5	0	5
Mississippi	3	1	2
Missouri	2	0	2
Montana	1	0	1
Nebraska	0	0	0
Nevada	6	0	6
New Hampshire	1	0	1
New Jersey	14	3	11
New Mexico	1	0	1
New York	13	1	12
New York City	23	3	20
North Carolina	9	1	8
North Dakota	1	0	1
Ohio	10	0	10
Oklahoma	4	0	4
Oregon	4	0	4
Pennsylvania	26	4	22
Rhode Island	2	0	2
South Carolina	6	1	5
South Dakota	0	0	0
Tennessee	16	3	13
Texas	32	8	24
Utah	3	3	0
Vermont	2	0	2
Virginia	13	1	12
Washington	7	1	6
West Virginia	3	0	3
Wisconsin	7	2	5
Wyoming	0	0	0
Puerto Rico	0	0	0
Virgin Islands	0	0	0
Total U.S.	385	51	334

* This column indicates how many of the 'TOTAL POSSIBLE RECORD MATCHES' resulted from situations where 'ALL' items provided on a user's request record matched with items on an NDI record.

** This column indicates how many of the 'TOTAL POSSIBLE RECORD MATCHES' resulted from situations where 'ONLY SOME' of the items provided on a user's request record matched with items on an NDI record.

NDI Retrieval Report [REPORT]

NOTE: If you decide to use computer software to assess the quality of NDI record matches or to manipulate the NDI output, use the record format provided in the "Combined File of Matching User and NDI Records" (file name = **COMBINED**).

The NDI Retrieval Report [**REPORT**] file is a PDF file intended only for you to print in **landscape** print format. Reduce margin widths if necessary.

The Retrieval Report is the primary output of an NDI file search. This report indicates which user records are involved in **possible** record matches with one or more NDI records. These possible matches are based on the seven matching criteria detailed at the beginning of Chapter 4. Major features of the **REPORT** file are:

1. Only your user records that generate **possible** record matches with NDI records (according to the seven matching criteria) are listed in the Retrieval Report. Your user records are listed in the same sequence in which you submitted them. User records that do not generate any possible NDI record matches are not included in the **REPORT** file. **If you submit a large volume of records, your REPORT file may only contain a partial listing of your matching results.**

2. The Retrieval Report first shows the information contained on the user record, followed by a listing of the NDI record or records that were identified as **possible** record matches.

3. More than one possible NDI record match may be generated for a particular user record. Several user records and their associated possible NDI record matches are listed in the sample Retrieval Report shown in Exhibit 6.

4. Information on each possible NDI record match includes state of death, death certificate number, and date of death. Next is indicated which data items on the NDI record agreed or disagreed with data items on your user record.

5. NDI record matches are listed in **ranked order** based solely on the number of data items that agree with those on the user record. The NDI record match containing the greatest number of matching data items is listed first. In addition, if all nine digits of the Social Security number agree, that NDI record would be listed first.

6. When ALL of the data items provided on your user record agree exactly with the related items on an NDI record, the NDI record match will be listed with an asterisk (*) in the left margin. This does not imply that the NDI record is a **true** or correct match, especially if you provided only a limited number of data items on the subject's record, or if the subject had a common name.

7. To properly assess your NDI results, it is important to become very familiar with the abbreviations and symbols used in the Retrieval Report. A description of the abbreviations and symbols is presented in Exhibit 6A. Geographic codes and abbreviations are listed in Exhibit 2B. Additional information regarding the assessment of your NDI results is provided in Chapter 4.

8. If you submit a large number of records, you may find even the partial printout of this entire file too cumbersome to use. You may need to use the **COMBINED** file described in the next section.

Exhibit 6

NDI Retrieval Report [REPORT]

All names and other data elements used in this exhibit are hypothetical.

NDI — NATIONAL DEATH INDEX RETRIEVAL REPORT — CDC

USER RECORD: (POSSIBLE MATCHES = 001) NDI Search NO : 2003-X456 CONTROL NO: 0046 USER DATA: 092698

POSSIBLE DECEDENT NAME	FATHER'S SURNAME	S S N	BIRTH DATE MO DAY YEAR	AGE	SEX	RACE	Marital Status	Res State	Birth State
RAYMOND ALDER	-	3 5 1 7 6 6 1 5 3	05 / 16 / 1948	47	M	-	-	CA	CA

POSSIBLE NDI RECORD MATCHES (IN RANKED ORDER)

STATE OF DEATH	CERTIFICATE NUMBER	DATE OF DEATH	FN	MI	LN	FATHER'S SURNAME	LN/FS	S S N	BIRTH MO	DATE DAY	+/- YEAR	AGE	SEX	RACE	Marital Status	Res State	Birth State	Status Code
*California	513478	06/15/1995	X	B	X	-	-	X X X X X X X X X	X	X	X	X	X	-	-	X	X	1

USER RECORD: (POSSIBLE MATCHES = 001) NDI APPL NO : Y3-X456 CONTROL NO: 0035 USER DATA: 091598

POSSIBLE DECEDENT NAME	FATHER'S SURNAME	S S N	BIRTH DATE MO DAY YEAR	AGE	SEX	RACE	Marital Status	Res State	Birth State
JANE C THOMPSON	SMITH	5 0 8 3 5 5 4 7 3	11 / 03 / 1960	42	F	-	-	NY	-

POSSIBLE NDI RECORD MATCHES (IN RANKED ORDER)

STATE OF DEATH	CERTIFICATE NUMBER	DATE OF DEATH	FN	MI	LN	FATHER'S SURNAME	LN/FS	S S N	BIRTH MO	DATE DAY	+/- YEAR	AGE	SEX	RACE	Marital Status	Res State	Birth State	Status Code
Alabama	010725	11/14/2002	X	X	X	X	-	X X X X X X X X X	X	X	X	X	X	-	-		-	1

USER RECORD: (POSSIBLE MATCHES = 004) NDI APPL NO : Y3-X456 CONTROL NO: 0156 USER DATA: 050988

POSSIBLE DECEDENT NAME	FATHER'S SURNAME	S S N	BIRTH DATE MO DAY YEAR	AGE	SEX	RACE	Marital Status	Res State	Birth State
ALBERT N BRADLEY	-	0 7 8 3 6 3 4 8 5	01 / 27 / 1953	40	M	-	-	NC	LA

POSSIBLE NDI RECORD MATCHES (IN RANKED ORDER)

STATE OF DEATH	CERTIFICATE NUMBER	DATE OF DEATH	FN	MI	LN	FATHER'S SURNAME	LN/FS	S S N	BIRTH MO	DATE DAY	+/- YEAR	AGE	SEX	RACE	Marital Status	Res State	Birth State	Status Code
North Carolina	069943	09/18/1993	X	?	X	-	-	X X X X X X X X X	X	X	X	X	X	-	-	X	X	1
Utah	130911	12/06/1999	X		N	-	-	- X X - - - - - -	X	X	-44		X	-	-		-	0
Maine	002151	01/23/2000	X	?	N	-	-	- x - x - - - - X	X		+01		X	-	-		-	0
California	173589	10/26/1985	X		N	-	-	- - - - - - X - -	X		+01		X	-	-		-	0

Retrieval Report (Abbreviations and Symbols)

Column heading abbreviations

Abbreviation	Definition
LN/FS	The **LAST NAME (LN)** supplied on the user record is compared with the **FATHER'S SURNAME (FS)** on the NDI record. This column is relevant only when SEX is female AND the user does not supply a FATHER'S SURNAME. A dash (-) appears in this column whenever this check is not performed; that is, whenever (a) sex is male or unknown or (b) father's surname is supplied on the user's record (for a direct comparison with the father's surname on the NDI record).
AGE	Age at death (actual or estimated)
YR	Year(s)
MO	Month(s)
WK	Week(s)
DY	Day(s)
HR	Hour(s)
MI	Minutes(s)
SEX	
M	Male
F	Female
MS	Marital status
S	Single
M	Married
W	Widowed
D	Divorced
RACE	
WH	White
BL	Black
IN	Indian
CH	Chinese
FI	Filipino
JA	Japanese
HA	Hawaiian
OA	Other Asian or Pacific Islander
ON	Other non-White
Res State	State of residence (see abbreviations in Exhibit 2B, Chapter 2)
Birth State	State of birth (see abbreviations in Exhibit 2B, Chapter 2)

Symbols used within report

Symbols	Definition
X	User's data item matches exactly with the comparable data item on the NDI record.
Blank	Data item supplied by the user DOES NOT match the data item on the NDI record.
?	Insufficient information on the NDI record; data item was provided on the user record.

Retrieval Report (Abbreviations and Symbols)—Con.

Symbols used within report

Symbols	Definition
- (dash)	General usage: Data item NOT provided on user record; see exceptions below.†
*	Asterisk in the left margin indicates that ALL data items provided on the user record match exactly with the related items on the NDI record.
N	Names match only New York State Identification and Intelligence System (NYSIIS) phonetic codes.
I	Only the first initials of the first names agreed on the user and NDI records. (The remaining characters of the first names on both records either did not agree or were not provided.)
B	Middle initials were NOT provided on either the user record or the NDI record (i.e., BOTH fields were blank). This occurrence is treated as a match on middle initial and ASSUMES that the individual(s) did not have a middle name. (A "blank" signifies that middle initials were provided but did not agree. An "X" signifies agreement on the middle initials. A dash (-) signifies that a middle initial existed only on the NDI record, while a "?" signifies that the middle initial existed only on the user record.)
A	The NDI record is an **ALIAS** record (also referred to as an "Also Known As" or "AKA" record). Some states will create and submit more than one NDI record for a decedent whenever the death certificate indicates that the decedent went by more than one first name and/or last name. The death certificate numbers on these NDI records will be the same, indicating only one death certificate on file in that state. (Beginning with deaths for 2010, NDI will also create two alias records if the main state death record is submitted with a dash (-) between two names in the last name field. Such state records, if listed as possible matches, will also have an "A" preceding the death certificate numbers.)
+01	The birth year on the NDI record is one year more than the birth year on the user record; for example, 1910 (NDI record) minus 1909 (user record) = +01. (Matching criteria 2 or 3 may have been satisfied; see **Chapter 4**.)
-01	The birth year on the NDI record is one year less than the birth year on the user record; for example, 1910 (NDI record) minus 1911 (user record) = -01. (Matching criteria 2 or 3 may have been satisfied; see **Chapter 4**.)
±02 through ±99	Difference between the years of birth on the NDI and user records. The four-digit birth year on the user record is always subtracted from the four-digit birth year on the NDI record. For example, 1905 (NDI record) minus 1920 (user) = -15 years.
>99	**Difference** between the **BIRTH YEARS** on the NDI record and the user record is greater than 99 years. For example, 2001 (NDI record) minus 1898 (user record) = 101 years or *greater than* 99 years.

†Exceptions for dashes:

- SSN: Specific digits of Social Security Number did not agree.
- SSN: When Social Security Number is not provided on user record, only one "-" is printed.
- Middle initial: No middle initial was provided on user record; however, a middle initial was provided on NDI record.
- LN/FS: No cross-check was performed between the LAST NAME (LN) on the user record and the FATHER'S SURNAME (FS) on the NDI record as described above.

Combined File of Matching User and NDI Records [COMBINED]

NOTE: This file contains the same information on each possible match as presented in the **Retrieval Report**.

1. The **COMBINED** file is intended for those NDI users who receive a large number of possible NDI record matches and would like to write computer programs to assist them in assessing the quality of the matches. This file (containing 178-character records) is not intended to be printed or read on a personal computer screen.

2. This file **combines** the information provided on a user record with information on the matching NDI record(s). Each **COMBINED** record on this file contains information from only one user record and only one matching NDI record. A separate **COMBINED** record is created for each NDI record match listed for a given user record in the NDI Retrieval Report. In such instances, all of the information from the user record is repeated at the beginning of each COMBINED RECORD.

3. The record layout for the **COMBINED** file is presented in Exhibit 7.

4. The **first 100 positions** of each **COMBINED** record contain the actual data from the user record. Refer to Chapter 2, Exhibit 2, for more detailed information about the code structure for this portion of the COMBINED RECORD.

5. **Positions 101–171** contain data related to the matching NDI record listed in the NDI Retrieval Report [**REPORT**] file. Refer to **Retrieval Report Abbreviations and Symbols** (Exhibit 6A) for more detailed information about the code structure for this portion of the **COMBINED** record.

6. **Positions 172–178** of each **COMBINED** record contain probabilistic scoring information on each match. Refer to Appendix A for information on probabilistic scoring.

7. The **COMBINED** record contains the state of death (data element 14) as well as the state of death **code** (data element 15) to permit users to assess agreements between the state of death code and the codes for the state of residence or state of birth (data elements 9 and 10, respectively).

8. The **COMBINED** record also indicates whether multiple NDI record matches occurred for a particular user record. Refer to **data element 32** (matching sequence) and **data element 33** (number of possible NDI record matches).

9. The **COMBINED** record contains military death certificate numbers in addition to state death certificate numbers. The military death certificate number is 1 letter + 5 digits. The state death certificate number is 6 digits.

Record Layout for Combined File [COMBINED]

*SPECIAL NOTE: This record layout is identical to the first 178 positions of the CAUSE file (see Exhibit 8). However, all of the NDI matching records listed in the COMBINED file do not appear in the CAUSE file (see Cause-of-Death File).

USER RECORD INFORMATION: Data elements 1–13, positions 1–100 (*Actual* information provided by NDI user)

	Data Elements	Positions	Code Structure
1.	Name of person in study group:		
	Last name	1–20	Alpha, left justified (see Exhibit 2A, Chapter 2)
	First name	21–35	Alpha, left justified (see Exhibit 2A, Chapter 2)
	Middle initial	36	Alpha or blank
2.	Social Security Number	37–45	Alpha/Numeric or blank
3.	Date of birth:		
	Month	46–47	01–12, 99, or blanks
	Day	48–49	01–31, 99, or blanks
	Year	50–53	1850–20xx or blanks
4.	Father's surname	54–71	Alpha, left justified; or blanks (see Exhibit 2A, Chapter 2)
5.	Age at death:		
	Unit	72	0–6, 9, or blanks (see Exhibit 2, element 5, Chapter 2)
	Number of units	73–74	00–98, 99, or blanks (see Exhibit 2, element 5, Chapter 2)
6.	Sex	75	1, 2, 9, M, F, or blank (see Exhibit 2, element 6, Chapter 2)
7.	Race	76	0–8, 9, or blank (see Exhibit 2, element 7, Chapter 2)
8.	Marital status	77	1–4, 9, or blank (see Exhibit 2, element 8, Chapter 2)
9.	State of residence	78–79	01–57, 59, 99, or blanks (see Exhibit 2B, Chapter 2)
10.	State of birth	80–81	01–57, 59, 99, or blanks (see Exhibit 2B, Chapter 2)
11.	Control or sequence number (optional)	82–91	Alpha/Numeric, left justified. (If this field is blank on user's record, NCHS assigns a sequence number.)
12.	Optional user data	92–97	Alpha/Numeric, left justified; or blanks. User may provide information or codes to assess each match (e.g., date of last contact or date of death; see Exhibit 2, element 12).
13.	Blank field	98–100	Blanks

Record Layout for Combined File [COMBINED]—Con.

NDI RECORD INFORMATION: Data elements 14–33, positions 101–178 (Information on matching NDI records)

	Data Elements	Positions	Code Structure
14.	State of death	101–112	Alpha, left justified
14A.	Year of death	113–116	Beginning with death year 1979 (4-digit year)
15.	State of death CODE	117–119	01–57, 59, left justified; 33C if New York City (see Exhibit 2B, Chapter 2)
16.	Alias indicator	120	"A" if an alias record; otherwise, blank
17.	Death certificate number	121–126	The state death certificate number is 6 digits. The military death certificate number is 1 letter + 5 digits. A = Army N = Navy F = Air Force M = Marine Corps C = Coast Guard
18.	Date of death:		
	Month	127–128	01–12
	Day	129–130	01–31
	Year (last 2 digits)	131–132	00–99
For items 19–31 (positions 133–165) below, see Exhibits 6 and 6A, Chapter 3			
19.	Name (NDI record):		
	First name	133–134	bX, bN, b?, lb, IN, bb (b = blank space)
	Middle initial	135	X, B, ?, -, b (b = blank space)
	Last name	136	X, N, ?, b (b = blank space)
20.	Father's surname (NDI record)	137	X, N, ?, -, b (b = blank space)
21.	Last name (on user record) compared with father's surname (on NDI record)	138	X, N, ?, -, b (b = blank space)
22.	Social Security Number (NDI record)	139–147	A string of 9 X's and/or dashes (-) to show agreement or disagreement on each digit; or one dash in position 143 (surrounded by blanks) to indicate that the SSN was not reported on user's record; or a question mark (?) in position 143 to indicate that the SSN was not reported only on the NDI record.
23.	Birth date (NDI record):		
	Month	148	X, ?, -, b (b = blank space)
	Day	149	X, ?, -, b (b = blank space)
	Year	150–152	Xbb, ?bb, -b (b = blank space) +01, -01, +02, -02, … +99, -99, >99
24.	Age at death (NDI record)	153	X, ?, -, b (b = blank space)
25.	Sex (NDI record)	154	X, ?, -, b (b = blank space)

Record Layout for Combined File [COMBINED]—Con.

	Data Elements	Positions	Code Structure
26.	Race (NDI record)	155	X, ?, -, b (b = blank space)
27.	Marital status (NDI record)	156	X, ?, -, b (b = blank space)
28.	State of residence (NDI record)	157	X, ?, -, b (b = blank space)
29.	State of birth (NDI record)	158	X, ?, -, b (b = blank space)
30.	Blank field	159–164	Blanks
31.	Exact match indicator	165	Asterisk (*) or blank (An "*" means all items provided on the user record match exactly with the corresponding items on the NDI record.)
32.	Matching sequence	166–168	001–050 This number indicates which multiple NDI record match is represented on this "User/NDI record" (e.g., **003** of 010 possible NDI record matches). **NOTE:** Virtually all of the NDI records presented in the **CAUSE** file will be the first ranked NDI record match (e.g., **001** of 010 possible matches). The **CAUSE** file may also include a few matches that were NOT ranked **first**; however, each of these matches had a probabilistic score high enough for the match to be assigned a status code of "1" (true match; assumed dead); see data element 36, position 178 below.
33.	Number of possible NDI record matches	169–171	001–050 This number indicates how many multiple NDI record matches were listed for a given user record (e.g., 003 of **010** possible NDI record matches; also refer to data element 32 above).
	PROBABILISTIC SCORING INFORMATION		(See Appendix A for more detailed information on probabilistic scoring.)
34.	PROBABILISTIC SCORE	172–176	
35.	CLASS CODE	177	1, 2, 3, 4, or 5
36.	STATUS CODE	178	0 = FALSE match (assumed alive) 1 = TRUE match (assumed dead)

Cause-of-Death File [CAUSE]

NOTE: DO NOT ACCESS THIS FILE FIRST. The NDI user must first access other NDI files to identify TRUE matches from all of the possible matches listed (see item 5). The CAUSE file does not contain all of the matches—it contains only those NDI matches (and their related cause-of-death codes) that were ranked first or had a high probabilistic score (see "Important Limitations" section).

If you selected true matches for which cause-of-death codes are not available on the CAUSE or PRTCAUSE files, contact NDI staff before attempting to purchase copies of death certificates. The staff may be able to provide you with the appropriate cause-of-death codes.

1. The NDI Plus **CAUSE** file does not appear in the NDI results for those users who requested only a **routine** NDI search (which provides only the state of death, date of death, and death certificate number for each possible match).

2. The **CAUSE** file contains 438 positions for each record, as described in Exhibit 8. The **CAUSE** file is not intended to be printed. It is a text file that can be used with your own software. Its primary use is to link the **true** matches identified in the **COMBINED** file with their cause-of-death codes in the **CAUSE** file.

3. The **CAUSE** file begins with the same 178-record format as the **COMBINED** file (positions 1–178). The remaining fields on the **CAUSE** file contain the coded causes of death (see positions 180–438). The order of the user records listed in the **CAUSE** file will be the same as the order in the **COMBINED** file.

4. The underlying and multiple cause-of-death codes reported for deaths occurring from 1979 through 1998 come from the *International Classification of Diseases, Ninth Revision* (ICD–9), while the cause-of-death codes for deaths occurring in 1999 and later come from the *International Classification of Diseases, 10th Revision* (ICD–10). All ICD–9 codes are three or four digits. All ICD–10 codes begin with an alpha character followed by two or three digits. NOTE: NDI results do not include decimals in the underlying and multiple cause-of-death codes. The decimal is implied between the third and fourth digits for ICD–9 codes and between the second and third digits for ICD–10 codes.

5. The file that contains your NDI results has a folder named "**CAUSE CODES**." **BEFORE OPENING THE FILES CONTAINING THE ICD–9 AND ICD–10 CODES, REFER TO APPENDIX B OF THIS USER'S GUIDE TO LEARN MORE ABOUT THE DIFFERENT TYPES OF CAUSE-OF-DEATH CODES. It is important that you understand the differences among the underlying cause-of-death codes, the entity-axis codes, and the record-axis codes. (The NDI results also contain a separate folder named "Read Me First.")**

6. Before using the **CAUSE** file, the NDI user should determine which of the possible NDI record matches are likely to be **true** matches and which are only **questionable** matches. This may be done by accessing either of the following two files, depending on the overall number of possible NDI record matches you have to assess:

   - NDI Retrieval Report [**REPORT**]
   - Combined File of Matching User and NDI Records[**COMBINED**]

7. Once the true and questionable NDI record matches have been determined, the user can link these records to the **CAUSE** file to obtain the associated coded causes of death. The linkage may be performed based on the following fields on both the **COMBINED** and the **CAUSE** files:

   - State of death code (positions 117–119)
   - Death certificate number (positions 121–126)
   - Year of death (positions 131–132)

8. Military causes of death are primary external causes. ICD does not provide for the level of specificity for external causes.

Important limitations

NDI is authorized to release only the coded causes of death for either of the following two types of NDI record matches:

1. **Matches that are ranked first in the NDI list of possible NDI record matches.** NDI record matches are ranked based on the number of data items that agree, with a record containing an agreement on Social Security number always being ranked first. If you use the **COMBINED** file to assess the NDI matches, refer to positions 166–168 to determine how the record was ranked. **NOTE: Many matches that are ranked first can still be false matches.**

2. **Any match (regardless of its rank) that has a high-enough probabilistic score to be assigned a Status Code of "1" (meaning true match; assumed dead)**. To determine whether an NDI record was assigned a Status Code of 1, refer to any of the following files:

   - NDI Retrieval Report [**REPORT**]—last printed column
   - Combined File of Matching User and NDI Records[**COMBINED**]— position 178
   - Cause-of-Death File[**CAUSE**]— position 178
   - Cause-of-Death Report [**PRTCAUSE**]— last printed column

NOTE: Most NDI record matches that are good enough to be assigned a Status Code of 1 will also tend to be ranked first in the list of possible matches. However, the Status Codes are intended only as a guide for the NDI user. The probabilistic scoring system gives fairly conservative results. Consequently, most of the records assigned a Status Code of 1 will actually be TRUE matches, while a number of additional true matches may be found among those records assigned the Status Code of 0 (meaning a false match; assumed alive). IT IS THE USER'S RESPONSIBILITY TO MAKE A FINAL DETERMINATION OF THE TRUE AND FALSE MATCHES. Refer to Appendix A for information on the assignment of probabilistic scores.

Other limitations

Illinois cause-of-death codes: Illinois state law prohibits the release of identifying cause-of-death information whenever selected causes-of-death appear on a death certificate. In such instances, the NDI Plus service is restricted from releasing any cause-of-death codes appearing on that record. This restriction pertains to only 1% of all the Illinois deaths occurring after 1986. For these records, the underlying cause field will contain "N/A" to indicate that the code is not available. The multiple-cause fields will be left blank. (NOTE: Whenever Illinois releases copies of such death certificates, they are required by law to remove all the causes of death that appear on the document.)

Unlinkable cause-of-death information: Because of minor differences each year in the cutoff dates before the creation of the NDI file and the NCHS Multiple Cause-of-Death files, a few instances may occur in which coded cause-of-death information is not available for the matching NDI records. For these records, the underlying cause field will contain "N/A" to indicate that the code is not available. The multiple cause-of-death fields will be left blank. This is especially true for those few **belated records** that arrive after a particular year of death has been closed.

1981 and 1982 multiple-cause data: With the exception of calendar years 1981 and 1982, all deaths occurring annually in the United States are processed. Although underlying cause data have been processed for all 54 registration areas, in 1981 and 1982 multiple-cause data were coded on only a 50% sample basis for deaths occurring in 19 registration areas. The remaining 35 registration areas had multiple-cause data processed on a 100% basis in 1981 and 1982. (The participating registration areas at the time were the 50 states, New York City, District of Columbia, Puerto Rico, and U.S. Virgin Islands.) The multiple-cause data for the following 19 registration areas were coded for only a 50% sample of their 1981 and 1982 deaths:

Alabama	Illinois	Oklahoma
Arizona	Indiana	Rhode Island
Arkansas	Kentucky	Tennessee
California	Montana	Texas
Florida	North Dakota	Washington
Georgia	Ohio	West Virginia
Idaho		

Record Layout for Cause File [CAUSE]

NOTE: Positions 1–178 are identical to "Exhibit 7: Record Layout for the **COMBINED** File."

USER RECORD INFORMATION: Data elements 1–13, positions 1–100 (*Actual* information provided by NDI user)

	Data Elements	Positions	Code Structure
1.	Name of person in study group:		
	Last name	1–20	Alpha, left justified (see Exhibit 2A, Chapter 2)
	First name	21–35	Alpha, left justified (see Exhibit 2A, Chapter 2)
	Middle initial	36	Alpha or blank
2.	Social Security Number	37–45	Alpha/Numeric or blank
3.	Date of birth:		
	Month	46–47	01–12, 99, or blanks
	Day	48–49	01–31, 99, or blanks
	Year	50–53	1850–20xx or blanks
4.	Father's surname	54–71	Alpha, left justified; or blanks (see Exhibit 2A, Chapter 2)
5.	Age at death:		
	Unit	72	0–6, 9, or blanks (see Exhibit 2, element 5, Chapter 2)
	Number of units	73–74	00–98, 99, or blanks (see Exhibit 2, element 5, Chapter 2)
6.	Sex	75	1, 2, 9, M, F, or blank (see Exhibit 2, element 6, Chapter 2)
7.	Race	76	0–8, 9, or blank (see Exhibit 2, element 7, Chapter 2)
8.	Marital status	77	1–4, 9, or blank (see Exhibit 2, element 8, Chapter 2)
9.	State of residence	78–79	01–57, 59, 99, or blanks (see Exhibit 2B, Chapter 2)
10.	State of birth	80–81	01–57, 59, 99, or blanks (see Exhibit 2B, Chapter 2)
11.	Control or sequence number (optional)	82–91	Alpha/Numeric, left justified. (If this field is blank on user's record, NCHS assigns a sequence number.)
12.	Optional user data	92–97	Alpha/Numeric, left justified; or blanks. User may provide information or codes to assess each match (e.g., date of last contact or date of death; see Exhibit 2, element 12).
13.	Blank field	98–100	Blanks

Record Layout for Cause File [CAUSE]—Con.

NDI RECORD INFORMATION: Data elements 14–33, positions 101–178 (Information on matching NDI records)

	Data Elements	Positions	Code Structure
14.	State of death	101–112	Alpha, left justified
14A.	Year of death	113–116	Beginning with death year 1979 (4-digit year)
15.	State of death CODE	117–119	01–57, 59, left justified; 33C if New York City (see Exhibit 2B, Chapter 2)
16.	Alias indicator	120	"A" if an alias record; otherwise, blank
17.	Death certificate number	121–126	The state death certificate number is 6 digits. The military death certificate number is 1 letter + 5 digits. A = Army N = Navy F = Air Force M = Marine Corps C = Coast Guard
18.	Date of death:		
	Month	127–128	01–12
	Day	129–130	01–31
	Year (last 2 digits)	131–132	00–99
For items 19–31 (positions 133–165) below, see Exhibit 6 and 6A, Chapter 3			
19.	Name (NDI record):		
	First name	133–134	bX, bN, b?, lb, IN, bb (b = blank space)
	Middle initial	135	X, B, ?, -, b (b = blank space)
	Last name	136	X, N, ?, b (b = blank space)
20.	Father's surname (NDI record)	137	X, N, ?, -, b (b = blank space)
21.	Last name (on user record) compared with father's surname (on NDI record)	138	X, N, ?, -, b (b = blank space)
22.	Social Security Number (NDI record)	139–147	A string of 9 Xs and/or dashes (-) to show agreement or disagreement on each digit; or one dash in position 143 (surrounded by blanks) to indicate that the SSN was not reported on user's record; or a question mark (?) in position 143 to indicate that the SSN was not reported only on the NDI record.
23.	Birth date (NDI record):		
	Month	148	X, ?, -, b (b = blank space)
	Day	149	X, ?, -, b (b = blank space)
	Year	150–152	Xbb, ?bb, -b (b = blank space) +01, -01, +02, -02, … +99, -99, >99
24.	Age at death (NDI record)	153	X, ?, -, b (b = blank space)
25.	Sex (NDI record)	154	X, ?, -, b (b = blank space)
26.	Race (NDI record)	155	X, ?, -, b (b = blank space)

Record Layout for Cause File [CAUSE]—Con.

	Data Elements	Positions	Code Structure
27.	Marital status (NDI record)	156	X, ?, -, b (b = blank space)
28.	State of residence (NDI record)	157	X, ?, -, b (b = blank space)
29.	State of birth (NDI record)	158	X, ?, -, b (b = blank space)
30.	Blank field	159–164	Blanks
31.	Exact match indicator	165	Asterisk (*) or blank (An "*" means all items provided on the user record match exactly with the corresponding items on the NDI record.)
32.	Matching sequence	166–168	001–050 This number indicates which multiple NDI record match is represented on this "User/NDI record" (e.g., **003** of 010 possible NDI record matches). **NOTE:** Virtually all of the NDI records presented in the CAUSE file will be the first ranked NDI record match (e.g., **001** of 010 possible matches). The CAUSE file may also include a few matches that were NOT ranked **first**; however, each of these matches had a probabilistic score high enough for the match to be assigned a status code of "1" (true match; assumed dead); see data element 36, position 178 below.
33.	Number of possible NDI record matches	169–171	001–050 This number indicates how many multiple NDI record matches were listed for a given user record (e.g., 003 of **010** possible NDI record matches; also refer to data element 32 above).
	PROBABILISTIC SCORING INFORMATION		(See Appendix A for more detailed information on probabilistic scoring.)
34.	PROBABILISTIC SCORE	172–176	
35.	CLASS CODE	177	1, 2, 3, 4, or 5
36.	STATUS CODE	178	0 = FALSE match (assumed alive) 1 = TRUE match (assumed dead)
37.	Blank field	179	Blank

Record Layout for Cause File [CAUSE]—Con.

CAUSE-OF-DEATH CODES: Data elements 38–83, positions 180–438
NOTE: ICD–9 codes are used for 1979–1998 deaths. ICD–10 codes are used for deaths beginning 1999 and begin with alpha characters. When cause-of-death codes are NOT available, positions 180–183 will contain N/A and positions 184–438 will be blank.

UNDERLYING CAUSE OF DEATH	180-194	These four fields contain the ICD code for the underlying cause-of-death and recodes of the ICD code into selected groups.

	Data Elements	Positions	Code Structure
38.	**Underlying Cause-of-Death Code**: ICD codes (9th or 10th Revisions) Refer to Appendix B for more information about the Underlying Cause-of-Death Code and the recodes assigned by NCHS.	180-183	**(1)** Sources: World Health Organization. Manual of the International Classification of Diseases, Injuries, and Causes of Death, based on recommendations of the Ninth Revision Conference, 1975. Geneva: World Health Organization. 1977. (2) World Health Organization. International Statistical Classification of Diseases and Related Health Problems, 10th Revision. Geneva: World Health Organization. 1992. (Refer to your *NDI Search Results* for files of all ICD codes.) **NOTE:** Whenever the ICD code has four positions, an implied decimal occurs after the third position. Position 183 is blank for those causes that do not have a fourth digit. Also note that the ICD–9 codes for *External Causes of Death* are actually E800–E999; however, for data processing purposes, NCHS never includes the letter "E" in positions 180–183. Any 800–999 codes in positions 180–183 are *external* cause codes. Do not confuse these external cause codes with the ICD–9 *Nature of Injury* codes (800–999), which can appear **only** in the *multiple* cause data fields for entity-axis conditions (positions 197–336) and for record-axis conditions (positions 339–438). *Nature of Injury* codes are never used for the underlying cause-of-death. (Refer to your NDI Search Results for a separate list of ICD–9 *Nature of Injury* Codes; ICD–10 *Nature of Injury* codes range from S00–T98.)
39.	282 ICD–9 or 358 ICD–10 Cause Recode	184–188	A recode of the ICD–9 or ICD–10 cause code into 282 or 358 groups for use in NCHS publications. (Refer to your NDI Search Results for a complete list of recodes and the causes included.) **Code ranges: ICD–9** (00100–35800), not inclusive; **ICD–10** (00100– 45600), not inclusive.

Record Layout for Cause File [CAUSE]—Con.

	Data Elements	Positions	Code Structure
40.	72 ICD–9 or 113 ICD–10 Cause Recode	189–191	A recode of the ICD–9 or ICD–10 cause code into 72 or 113 groups for use in NCHS publications. (Refer to your NDI Search Results for a complete list of these recodes.) **Code ranges: ICD–9** (010–840), not inclusive; **ICD–10** (001–135), not inclusive.
41.	61 ICD–9 or 130 ICD–10 Infant Cause Recode	192–194	A recode of the ICD–9 or ICD–10 cause code into 61 or 130 groups for use in NCHS publications. Note that while records for all age groups will have a recode in this field (positions 192–194), the field is only applicable when compiling statistics on infants under age 1 year. **Code ranges: ICD–9** (010–680), not inclusive; **ICD–10** (001–158), not inclusive.
MULTIPLE CAUSES: ENTITY-AXIS CONDITIONS			
	ENTITY-AXIS CONDITIONS Refer to Appendix B for more information on entity-axis codes (or conditions).	197–336	Space has been provided for a maximum of 20 conditions. Each condition takes 7 positions in the record. Records that do not have 20 conditions are blank in the unused positions or area. Position 1: Part/line number on certificate 1 ..... Part I, line 1 (a) 2 ..... Part I, line 2 (b) 3 ..... Part I, line 3 (c) 4 ..... Part I, line 4 (d) 5 ..... Part I, line 5 (e) 6 ..... Part II Position 2: Sequence of condition within part/line code range: 1–7 Position 3–6: Condition code (See ICD–9 or ICD–10 codes in your NDI Search Results.) **NOTE:** Whenever a four-position code occurs, an implied decimal always occurs after the third position. Position 7: Nature of Injury flag (only for ICD–9 codes) 1 ..... Indicates that the code in positions 3–6 is a Nature of Injury ICD–9 code 0 .... All other codes
43.	1st Condition	197–203	
44.	2nd Condition	204–210	
45.	3rd Condition	211–217	
46.	4th Condition	218–224	
47.	5th Condition	225–231	
48.	6th Condition	232–238	
49.	7th Condition	239–245	
50.	8th Condition	246–252	
51.	9th Condition	253–259	

Record Layout for Cause File [CAUSE]—Con.

	Data Elements	Positions	Code Structure
52.	10th Condition	260–266	
53.	11th Condition	267–273	
54.	12th Condition	274–280	
55.	13th Condition	281–287	
56.	14th Condition	288–294	
57.	15th Condition	295–301	
58.	16th Condition	302–308	
59.	17th Condition	309–315	
60.	18th Condition	316–322	
61.	19th Condition	323–329	
62.	20th Condition	330–336	
MULTIPLE CAUSES: RECORD-AXIS CONDITIONS			
63.	Number of Record Axis Conditions	337–338	Code range: 00–20
	RECORD-AXIS CONDITIONS Refer to Appendix B for more information on record-axis codes (or conditions).		Space has been provided for a maximum of 20 conditions. Each condition takes 5 positions in the record. Records that do not have 20 conditions are blank in the unused positions or area. Positions 1–4: Condition Code (see ICD–9 or ICD–10 codes in your NDI Search Results.) **NOTE:** Whenever a four-position code occurs, an implied decimal always occurs after the third position. Position 5: Nature of Injury Flag (only for ICD–9 codes) 1 ..... Indicates that the code in position 1–4 is a Nature of Injury ICD–9 code. 0 ..... All other codes
64.	1st Condition	339–343	
65.	2nd Condition	344–348	
66.	3rd Condition	349–353	
67.	4th Condition	354–358	
68.	5th Condition	359–363	
69.	6th Condition	364–368	
70.	7th Condition	369–373	
71.	8th Condition	374–378	
72.	9th Condition	379–383	
73.	10th Condition	384–388	
74.	11th Condition	389–393	
75.	12th Condition	394–398	
76.	13th Condition	399–403	
77.	14th Condition	404–408	

Record Layout for Cause File [CAUSE]—Con.

	Data Elements	Positions	Code Structure
78.	15th Condition	409–413	
79.	16th Condition	414–418	
80.	17th Condition	419–423	
81.	18th Condition	424–428	
82.	19th Condition	429–433	
83.	20th Condition	434–438	

Cause-of-Death Report [PRTCAUSE]

The **PRTCAUSE** file represents the PRINTCAUSE file.

The **PRTCAUSE** file contains the **Cause-of-Death *Report*** that is formatted solely to generate an easy-to-read printout of the same coded cause-of-death included in the **Cause-of-Death *File*** (**CAUSE** file). An example of a printed page of this report appears in Exhibit 9. The cause-of-death codes used in Exhibit 9 are intended only to show the relative positions and lengths of the cause-of-death codes that would appear in an actual report.

The **PRTCAUSE** file is intended primarily for NDI users who:

3. Submit a small number of records for an NDI Plus search. (**NOTE: If you submitted a large volume of records, your PRTCAUSE file may contain only a partial listing of your search results.**)
4. Expect to identify only a small number of deaths.
5. Do not plan to write computer programs to link the coded causes of death from the CAUSE file with other data on the study subjects.
6. Want to quickly inspect the coded causes of death associated with **true** NDI record matches.

As discussed with the **CAUSE** file in the previous section, the **PRTCAUSE** file also lists only those NDI record matches that were either ranked first among the possible matches or received probabilistic scores high enough to be assumed to be TRUE matches. Based on the probabilistic score, each record is assigned a Status Code. The Status Codes can be found in the last column of the printed PRTCAUSE file. Each NDI record match listed will have either a status code of "1" (meaning TRUE match; assumed dead) or a status code of "0" (meaning FALSE match; assumed alive).

NOTE: Most NDI record matches that are good enough to be assigned a Status Code of 1 will also tend to be ranked first in the list of possible matches. However, the Status Codes are intended only as a guide for the NDI user. The probabilistic scoring system gives fairly conservative results. Consequently, most of the records assigned a Status Code of 1 will actually be TRUE matches, while a number of additional TRUE matches may be found among those records assigned the Status Code of 0 (meaning a FALSE match; assumed alive). IT IS THE USER'S RESPONSIBILITY TO MAKE A FINAL DETERMINATION OF THE TRUE AND FALSE MATCHES. Refer to Appendix A for information on the assignment of probabilistic scores.

The order in which possible NDI record matches are listed in the printed Cause-of-Death Report is as follows:

1. The matches are first grouped BY STATE (in alphabetical order; however, territories are listed last).
2. The matches are then listed BY YEAR OF DEATH (in ascending order).
3. Within a given year, the matches are listed BY DEATH CERTIFICATE NUMBER (in ascending order).

Refer to your "NDI Search Results" file for lists of all ICD cause-of-death codes. To better understand the various cause-of-death codes and conditions, ***you are strongly encouraged to review Appendix B of this Users Guide first.*** Note that ICD–9 codes are used for deaths occurring through 1998, and ICD–10 codes are used from 1999 onward. Furthermore, all ICD–10 codes start with an alpha character while ICD–9 codes are composed only of numbers.

Underlying cause of death: The following information about the codes for underlying cause-of-death and multiple causes-of-death is repeated from item 38 of the CAUSE file in the previous section:

NOTE: Whenever the ICD code has four positions, a decimal is implied after the third position. The fourth position is blank for those causes that do not have a 4th digit. Also note that the ICD–9 codes for External Causes-of-Death are actually E800–E999; however, for data processing purposes, NCHS never includes the letter "E" in the first position. Any 800–999 codes listed in the underlying cause-of-death box in Exhibit 9 are **external** cause codes. Do not confuse these external cause codes with the ICD–9 Nature of Injury codes (800–999), which can only appear in the multiple-cause rows for entity-axis conditions and for record-axis conditions. Nature of Injury codes are **never** used for the underlying cause-of-death. (Refer to your "NDI Search Results" file for a separate list of ICD–9 Nature of Injury codes. For ICD–10, the Nature of Injury codes and the External Causes of Death codes do not overlap. ICD–10 Nature of Injury codes range from S00 through T98, and External Causes of Death codes range from U01 through Y89.)

Entity-axis conditions: Each of these codes includes a prefix that indicates the relative order in which the condition was listed in the cause-of-death section of a death certificate. These prefixes are in positions 1–2. Positions 3–6 are the actual conditions—with an implied decimal after the 5th position. The 7th position is the Nature of Injury flag (only for ICD–9 codes). These various positions are further described as follows:

Space has been provided for a maximum of 20 conditions. Each condition takes 7 positions in the record. Records that do not have 20 conditions are blank in the unused positions or area.

Position 1: Part/line number on certificate

1 ..... Part I, line 1 (a)

2 ..... Part I, line 2 (b)

3 ..... Part I, line 3 (c)

4 ..... Part I, line 4 (d)

5 ..... Part I, line 5 (e)

6 ..... Part II

Position 2: Sequence of condition within part/line—Code range: 1–7

Positions 3-6: Condition code (See files containing ICD–9 and ICD–10 codes in your "NDI Search Results" file.) NOTE: Whenever a 4-position condition code occurs, a decimal is always implied after the 3rd position of that condition code.

Position 7: Nature of Injury flag (**only for ICD–9 codes**)

1 ..... Indicates that the code in positions 3–6 is a Nature of Injury ICD–9 code

0 ..... All other codes

Record-axis conditions: Whenever a 4-position code occurs, a decimal is always implied after the 3rd position. Record axis codes and conditions are further described as follows:

Space has been provided for a maximum of 20 conditions. Each condition takes five positions in the record. Records that do not have 20 conditions are blank in the unused positions or area.

Positions 1–4: Condition Code (See files containing ICD–9 and ICD–10 codes in your "NDI Search Results.")

Position 5: Nature of Injury flag (**only for ICD–9 codes**)

1 ..... Indicates that the code in position 1–4 is a Nature of Injury ICD–9 code

0 ..... All other codes

NOTE: An explanation of entity and record-axis conditions is described in Appendix B.

Exhibit 9

Cause-of-Death Report [PRTCAUSE]

All names and other data elements used in this exhibit are hypothetical.

NDI **NATIONAL DEATH INDEX RETRIEVAL REPORT** CDC

Death Year	State Code	State Name	Certificate No	Control No	Last Name (As Sent)	Underlying Cause Code	282 or 358 Cause Recode	72 or 113 Cause Recode	61 or 130 Cause Recode	Status Code
2002	01	ALABAMA	019137	0046	ALDER	X31	423	123		1

ENTITY AXIS CODES: 11I469 21T68 22X31

RECORD AXIS CODES: I469 T68 X31

Death Year	State Code	State Name	Certificate No	Control No	Last Name (As Sent)	Underlying Cause Code	282 or 358 Cause Recode	72 or 113 Cause Recode	61 or 130 Cause Recode	Status Code
2002	01	ALABAMA	035414	0035	JONES	C499	103	043		1

ENTITY AXIS CODES: 11J969 21C499

RECORD AXIS CODES: C499 J969

Death Year	State Code	State Name	Certificate No	Control No	Last Name (As Sent)	Underlying Cause Code	282 or 358 Cause Recode	72 or 113 Cause Recode	61 or 130 Cause Recode	Status Code
2002	04	ARKANSAS	014670	0146	MURPHY	K703	298	094		0

ENTITY AXIS CODES: 11K922 21F102 61K703 62I850

RECORD AXIS CODES: F102 I850 K703 K922

Death Year	State Code	State Name	Certificate No	Control No	Last Name (As Sent)	Underlying Cause Code	282 or 358 Cause Recode	72 or 113 Cause Recode	61 or 130 Cause Recode	Status Code
1992	05	CALIFORNIA	013392	0087	WILSON	3310	15300	780		1

ENTITY AXIS CODES: 1103800 21485 0 3133100

RECORD AXIS CODES: 03800 33100 485 0

Matching User Records [MATCH], Nonmatching User Records [NOMATCH], and Rejected User Records [REJECTS]

1. These three files contain only the user record information you provided. These records are in the same record format as the records you submitted to NCHS.

2. The user records are grouped as follows:

   a. Matching user records (**MATCH** file): Contains user records involved in a possible match with at least one NDI record.

   b. Nonmatching user records (**NOMATCH** file): Contains user records not involved in any match with NDI records.

   c. Rejected user records (**REJECTS** file): Contains user records that failed to satisfy the basic criteria of the NDI edit program and were thus rejected **prior to** the search of the NDI file.

3. These three files are generated strictly for your information. They potentially can be used as follows:

   a. **MATCH** file: After you determine which of the MATCHING USER RECORDS were not involved in **true** matches, you can delete these records from the **MATCH** file and add them to the **NOMATCH** file.

   b. **NOMATCH** file: This file may be saved and then resubmitted to NCHS for repeat NDI file searches as subsequent years of death become available.

   c. **REJECTS** file: Each record on this file was REJECTED because it did not contain at least a first and last name and either a Social Security number or a month and year of birth. Note, however, that if a user record does not contain a first and last name, it will not be rejected as long as it at least has a Social Security number, date of birth, and sex. If the missing information on some or all of these user records can be obtained and added to the file, the records could be resubmitted for an NDI file search at no extra charge.

4. The record format and code structure for the user records on these three files are presented in Exhibit 2, Chapter 2.

NOTE: A copy of the U.S. Standard Certificate of Death appears in Exhibit 10 with instructions for completing the cause-of-death sections appearing in Exhibit 10A.

U.S. Standard Certificate of Death

LOCAL FILE NO. STATE FILE NO.

NAME OF DECEDENT
For use by physician or institution

To Be Completed/ Verified By: FUNERAL DIRECTOR:

1. DECEDENT'S LEGAL NAME (Include AKA's if any) (First, Middle, Last)
2. SEX
3. SOCIAL SECURITY NUMBER
4a. AGE-Last Birthday (Years)
4b. UNDER 1 YEAR: Months, Days
4c. UNDER 1 DAY: Hours, Minutes
5. DATE OF BIRTH (Mo/Day/Yr)
6. BIRTHPLACE (City and State or Foreign Country)
7a. RESIDENCE-STATE
7b. COUNTY
7c. CITY OR TOWN
7d. STREET AND NUMBER
7e. APT. NO.
7f. ZIP CODE
7g. INSIDE CITY LIMITS? ☐ Yes ☐ No
8. EVER IN US ARMED FORCES? ☐ Yes ☐ No
9. MARITAL STATUS AT TIME OF DEATH ☐ Married ☐ Married, but separated ☐ Widowed ☐ Divorced ☐ Never Married ☐ Unknown
10. SURVIVING SPOUSE'S NAME (If wife, give name prior to first marriage)
11. FATHER'S NAME (First, Middle, Last)
12. MOTHER'S NAME PRIOR TO FIRST MARRIAGE (First, Middle, Last)
13a. INFORMANT'S NAME
13b. RELATIONSHIP TO DECEDENT
13c. MAILING ADDRESS (Street and Number, City, State, Zip Code)
14. PLACE OF DEATH (Check only one: see instructions)
IF DEATH OCCURRED IN A HOSPITAL: ☐ Inpatient ☐ Emergency Room/Outpatient ☐ Dead on Arrival
IF DEATH OCCURRED SOMEWHERE OTHER THAN A HOSPITAL: ☐ Hospice facility ☐ Nursing home/Long term care facility ☐ Decedent's home ☐ Other (Specify):
15. FACILITY NAME (If not institution, give street & number)
16. CITY OR TOWN , STATE, AND ZIP CODE
17. COUNTY OF DEATH
18. METHOD OF DISPOSITION: ☐ Burial ☐ Cremation ☐ Donation ☐ Entombment ☐ Removal from State ☐ Other (Specify):
19. PLACE OF DISPOSITION (Name of cemetery, crematory, other place)
20. LOCATION-CITY, TOWN, AND STATE
21. NAME AND COMPLETE ADDRESS OF FUNERAL FACILITY
22. SIGNATURE OF FUNERAL SERVICE LICENSEE OR OTHER AGENT
23. LICENSE NUMBER (Of Licensee)

To Be Completed By: MEDICAL CERTIFIER

ITEMS 24-28 MUST BE COMPLETED BY PERSON WHO PRONOUNCES OR CERTIFIES DEATH

24. DATE PRONOUNCED DEAD (Mo/Day/Yr)
25. TIME PRONOUNCED DEAD
26. SIGNATURE OF PERSON PRONOUNCING DEATH (Only when applicable)
27. LICENSE NUMBER
28. DATE SIGNED (Mo/Day/Yr)
29. ACTUAL OR PRESUMED DATE OF DEATH (Mo/Day/Yr) (Spell Month)
30. ACTUAL OR PRESUMED TIME OF DEATH
31. WAS MEDICAL EXAMINER OR CORONER CONTACTED? ☐ Yes ☐ No

CAUSE OF DEATH (See instructions and examples)

32. **PART I.** Enter the chain of events--diseases, injuries, or complications--that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary.

Approximate interval: Onset to death

IMMEDIATE CAUSE (Final disease or condition resulting in death) --------> a.
Due to (or as a consequence of):
Sequentially list conditions, if any, leading to the cause listed on line a. Enter the **UNDERLYING CAUSE** (disease or injury that initiated the events resulting in death) **LAST**
b.
Due to (or as a consequence of):
c.
Due to (or as a consequence of):
d.

PART II. Enter other significant conditions contributing to death but not resulting in the underlying cause given in PART I

33. WAS AN AUTOPSY PERFORMED? ☐ Yes ☐ No
34. WERE AUTOPSY FINDINGS AVAILABLE TO COMPLETE THE CAUSE OF DEATH? ☐ Yes ☐ No
35. DID TOBACCO USE CONTRIBUTE TO DEATH? ☐ Yes ☐ Probably ☐ No ☐ Unknown
36. IF FEMALE: ☐ Not pregnant within past year ☐ Pregnant at time of death ☐ Not pregnant, but pregnant within 42 days of death ☐ Not pregnant, but pregnant 43 days to 1 year before death ☐ Unknown if pregnant within the past year
37. MANNER OF DEATH ☐ Natural ☐ Homicide ☐ Accident ☐ Pending Investigation ☐ Suicide ☐ Could not be determined
38. DATE OF INJURY (Mo/Day/Yr) (Spell Month)
39. TIME OF INJURY
40. PLACE OF INJURY (e.g., Decedent's home; construction site; restaurant; wooded area)
41. INJURY AT WORK? ☐ Yes ☐ No
42. LOCATION OF INJURY: State: City or Town: Street & Number: Apartment No.: Zip Code:
43. DESCRIBE HOW INJURY OCCURRED:
44. IF TRANSPORTATION INJURY, SPECIFY: ☐ Driver/Operator ☐ Passenger ☐ Pedestrian ☐ Other (Specify)
45. CERTIFIER (Check only one):
☐ Certifying physician-To the best of my knowledge, death occurred due to the cause(s) and manner stated.
☐ Pronouncing & Certifying physician-To the best of my knowledge, death occurred at the time, date, and place, and due to the cause(s) and manner stated.
☐ Medical Examiner/Coroner-On the basis of examination, and/or investigation, in my opinion, death occurred at the time, date, and place, and due to the cause(s) and manner stated.
Signature of certifier:
46. NAME, ADDRESS, AND ZIP CODE OF PERSON COMPLETING CAUSE OF DEATH (Item 32)
47. TITLE OF CERTIFIER
48. LICENSE NUMBER
49. DATE CERTIFIED (Mo/Day/Yr)
50. **FOR REGISTRAR ONLY-** DATE FILED (Mo/Day/Yr)

To Be Completed By: FUNERAL DIRECTOR

51. DECEDENT'S EDUCATION-Check the box that best describes the highest degree or level of school completed at the time of death.
☐ 8th grade or less
☐ 9th - 12th grade; no diploma
☐ High school graduate or GED completed
☐ Some college credit, but no degree
☐ Associate degree (e.g., AA, AS)
☐ Bachelor's degree (e.g., BA, AB, BS)
☐ Master's degree (e.g., MA, MS, MEng, MEd, MSW, MBA)
☐ Doctorate (e.g., PhD, EdD) or Professional degree (e.g., MD, DDS, DVM, LLB, JD)
52. DECEDENT OF HISPANIC ORIGIN? Check the box that best describes whether the decedent is Spanish/Hispanic/Latino. Check the "No" box if decedent is not Spanish/Hispanic/Latino.
☐ No, not Spanish/Hispanic/Latino
☐ Yes, Mexican, Mexican American, Chicano
☐ Yes, Puerto Rican
☐ Yes, Cuban
☐ Yes, other Spanish/Hispanic/Latino (Specify)
53. DECEDENT'S RACE (Check one or more races to indicate what the decedent considered himself or herself to be)
☐ White
☐ Black or African American
☐ American Indian or Alaska Native (Name of the enrolled or principal tribe)
☐ Asian Indian
☐ Chinese
☐ Filipino
☐ Japanese
☐ Korean
☐ Vietnamese
☐ Other Asian (Specify)
☐ Native Hawaiian
☐ Guamanian or Chamorro
☐ Samoan
☐ Other Pacific Islander (Specify)
☐ Other (Specify)
54. DECEDENT'S USUAL OCCUPATION (Indicate type of work done during most of working life. DO NOT USE RETIRED).
55. KIND OF BUSINESS/INDUSTRY

REV. 11/2003

Cause of Death—Background, Examples, and Common Problems

Accurate cause of death information is important
•to the public health community in evaluating and improving the health of all citizens, and
•often to the family, now and in the future, and to the person settling the decedent's estate.

The cause-of-death section consists of two parts. **Part I** is for reporting a chain of events leading directly to death, with the **immediate cause** of death (the final disease, injury, or complication directly causing death) on line a and the **underlying cause** of death (the disease or injury that initiated the chain of events that led directly and inevitably to death) on the lowest used line. **Part II** is for reporting all other significant diseases, conditions, or injuries that contributed to death but which did not result in the underlying cause of death given in **Part I**. **The cause-of-death information should be YOUR best medical OPINION.** A condition can be listed as "probable" even if it has not been definitively diagnosed.

Examples of properly completed medical certifications

CAUSE OF DEATH (See instructions and examples)

32. **PART I.** Enter the chain of events--diseases, injuries, or complications--that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary.

		Approximate interval: Onset to death
IMMEDIATE CAUSE (Final disease or condition ---------> resulting in death)	a. Rupture of myocardium Due to (or as a consequence of):	Minutes
Sequentially list conditions, if any, leading to the cause listed on line a. Enter the **UNDERLYING CAUSE** (disease or injury that initiated the events resulting in death) **LAST**	b. Acute myocardial infarction Due to (or as a consequence of):	6 days
	c. Coronary artery thrombosis Due to (or as a consequence of):	5 years
	d. Atherosclerotic coronary artery disease	7 years

PART II. Enter other significant conditions contributing to death but not resulting in the underlying cause given in PART I

Diabetes, Chronic obstructive pulmonary disease, smoking

33. WAS AN AUTOPSY PERFORMED? ■ Yes ☐ No

34. WERE AUTOPSY FINDINGS AVAILABLE TO COMPLETE THE CAUSE OF DEATH? ■ Yes ☐ No

35. DID TOBACCO USE CONTRIBUTE TO DEATH?
■ Yes ☐ Probably
☐ No ☐ Unknown

36. IF FEMALE:
■ Not pregnant within past year
☐ Pregnant at time of death
☐ Not pregnant, but pregnant within 42 days of death
☐ Not pregnant, but pregnant 43 days to 1 year before death
☐ Unknown if pregnant within the past year

37. MANNER OF DEATH
■ Natural ☐ Homicide
☐ Accident ☐ Pending Investigation
☐ Suicide ☐ Could not be determined

CAUSE OF DEATH (See instructions and examples)

32. **PART I.** Enter the chain of events--diseases, injuries, or complications--that directly caused the death. DO NOT enter terminal events such as cardiac arrest, respiratory arrest, or ventricular fibrillation without showing the etiology. DO NOT ABBREVIATE. Enter only one cause on a line. Add additional lines if necessary.

		Approximate interval: Onset to death
IMMEDIATE CAUSE (Final disease or condition ---------> resulting in death)	a. Aspiration pneumonia Due to (or as a consequence of):	2 Days
Sequentially list conditions, if any, leading to the cause listed on line a. Enter the **UNDERLYING CAUSE** (disease or injury that initiated the events resulting in death) **LAST**	b. Complications of coma Due to (or as a consequence of):	7 weeks
	c. Blunt force injuries Due to (or as a consequence of):	7 weeks
	d. Motor vehicle accident	7 weeks

PART II. Enter other significant conditions contributing to death but not resulting in the underlying cause given in PART I

33. WAS AN AUTOPSY PERFORMED? ■ Yes ☐ No

34. WERE AUTOPSY FINDINGS AVAILABLE TO COMPLETE THE CAUSE OF DEATH? ■ Yes ☐ No

35. DID TOBACCO USE CONTRIBUTE TO DEATH?
☐ Yes ☐ Probably
■ No ☐ Unknown

36. IF FEMALE:
☐ Not pregnant within past year
☐ Pregnant at time of death
☐ Not pregnant, but pregnant within 42 days of death
☐ Not pregnant, but pregnant 43 days to 1 year before death
☐ Unknown if pregnant within the past year

37. MANNER OF DEATH
☐ Natural ☐ Homicide
■ Accident ☐ Pending Investigation
☐ Suicide ☐ Could not be determined

38. DATE OF INJURY (Mo/Day/Yr) (Spell Month)
August 15, 2003

39. TIME OF INJURY
Approx. 2320

40. PLACE OF INJURY (e.g., Decedent's home; construction site; restaurant; wooded area)
road side near state highway

41. INJURY AT WORK?
☐ Yes ■ No

42. LOCATION OF INJURY: State: Missouri City or Town: near Alexandria
Street & Number: mile marker 17 on state route 46a Apartment No.: Zip Code:

43. DESCRIBE HOW INJURY OCCURRED:
Decedent driver of van, ran off road into tree

44. IF TRANSPORTATION INJURY, SPECIFY:
■ Driver/Operator
☐ Passenger
☐ Pedestrian
☐ Other (Specify)

Common problems in death certification

The **elderly decedent** should have a clear and distinct etiological sequence for cause of death, if possible. Terms such as senescence, infirmity, old age, and advanced age have little value for public health or medical research. Age is recorded elsewhere on the certificate. When a number of conditions resulted in death, the physician should choose the single sequence that, in his or her opinion, best describes the process leading to death, and place any other pertinent conditions in Part II. If after careful consideration the physician cannot determine a sequence that ends in death, then the medical examiner or coroner should be consulted about conducting an investigation or providing assistance in completing the cause of death.

The **infant decedent** should have a clear and distinct etiological sequence for cause of death, if possible. "Prematurity" should not be entered without explaining the etiology of prematurity. Maternal conditions may have initiated or affected the sequence that resulted in infant death, and such maternal causes should be reported in addition to the infant causes on the infant's death certificate (e.g., Hyaline membrane disease **due to** prematurity, 28 weeks **due to** placental abruption **due to** blunt trauma to mother's abdomen).

When **SIDS** is suspected, a complete investigation should be conducted, typically by a medical examiner or coroner. If the infant is under 1 year of age, no cause of death is determined after scene investigation, clinical history is reviewed, and a complete autopsy is performed, then the death can be reported as Sudden Infant Death Syndrome.

When processes such as the following are reported, additional information about the etiology should be reported:

Abscess
Abdominal hemorrhage
Adhesions
Adult respiratory distress syndrome
Acute myocardial infarction
Altered mental status
Anemia
Anoxia
Anoxic encephalopathy
Arrhythmia
Ascites
Aspiration
Atrial fibrillation
Bacteremia
Bedridden
Biliary obstruction
Bowel obstruction
Brain injury
Brain stem herniation
Carcinogenesis
Carcinomatosis
Cardiac arrest
Cardiac dysrhythmia
Cardiomyopathy
Cardiopulmonary arrest
Cellulitis
Cerebral edema
Cerebrovascular accident
Cerebellar tonsillar herniation
Chronic bedridden state
Cirrhosis
Coagulopathy
Compression fracture
Congestive heart failure
Convulsions
Decubiti
Dehydration
Dementia (when not otherwise specified)
Diarrhea
Disseminated intra vascular coagulopathy
Dysrhythmia
End-stage liver disease
End-stage renal disease
Epidural hematoma
Exsanguination
Failure to thrive
Fracture
Gangrene
Gastrointestinal hemorrhage
Heart failure
Hemothorax
Hepatic failure
Hepatitis
Hepatorenal syndrome
Hyperglycemia
Hyperkalemia
Hypovolemic shock
Hyponatremia
Hypotension
Immunosuppression
Increased intra cranial pressure
Intra cranial hemorrhage
Malnutrition
Metabolic encephalopathy
Multi-organ failure
Multi-system organ failure
Myocardial infarction
Necrotizing soft-tissue infection
Old age
Open (or closed) head injury
Paralysis
Pancytopenia
Perforated gallbladder
Peritonitis
Pleural effusions
Pneumonia
Pulmonary arrest
Pulmonary edema
Pulmonary embolism
Pulmonary insufficiency
Renal failure
Respiratory arrest
Seizures
Sepsis
Septic shock
Shock
Starvation
Subdural hematoma
Subarachnoid hemorrhage
Sudden death
Thrombocytopenia
Uncal herniation
Urinary tract infection
Ventricular fibrillation
Ventricular tachycardia
Volume depletion

If the certifier is unable to determine the etiology of a process such as those shown above, the process must be qualified as being of an unknown, undetermined, probable, presumed, or unspecified etiology so it is clear that a distinct etiology was not inadvertently or carelessly omitted.

The following conditions and types of death might seem to be specific or natural but when the medical history is examined further may be found to be complications of an injury or poisoning (possibly occurring long ago). Such cases should be reported to the medical examiner/coroner.

Asphyxia
Bolus
Choking
Drug or alcohol overdose/drug or alcohol abuse
Epidural hematoma
Exsanguination
Fall
Fracture
Hip fracture
Hyperthermia
Hypothermia
Open reduction of fracture
Pulmonary emboli
Seizure disorder
Sepsis
Subarachnoid hemorrhage
Subdural hematoma
Surgery
Thermal burns/chemical burns

REV. 11/2003

Assessing NDI Output

Review Your NDI Edit Results First

Before you begin assessing your NDI matches, it is important to carefully review the file named **EDITS** that is included in your NDI results. (Refer to the "User File Edit Results" section in Chapter 3 for more information on the **EDITS** file.)

Your study subjects' records are edited before they are searched against the NDI file. The NDI edit program determines whether your data file conforms to NCHS data file specifications, file format requirements, and coding instructions. If your data file cannot be read or does not satisfy the basic edit program requirements, your file will be returned to you for revision and resubmission at no extra charge.

The edit program rejects any user records not containing **at least one** of the following combinations of data items:

1. FIRST and LAST NAME and SOCIAL SECURITY NUMBER
2. FIRST and LAST NAME and MONTH and YEAR OF BIRTH
3. SOCIAL SECURITY NUMBER and DATE OF BIRTH and SEX

These combinations of items are part of the NDI Edit requirements that determine whether a user's record has sufficient data to qualify for a search. These combinations should not be confused with the following seven NDI matching criteria.

User records rejected by the edit program are not eligible for inclusion in the search of the NDI file. The edit results provide an error listing of each rejected user record with the reason for the rejection. If you have sufficient data to update the rejected records, you may correct and resubmit such records at no charge.

Furthermore, the edit results also include summary statistics on the accuracy and completeness of each NDI data set item you provided. If you notice that one or more of the data items you submitted are missing, out of range, or have an unexpectedly low completeness percentage, consider correcting those items and resubmitting a corrected file for another NDI search at no charge.

Matching Criteria

The NDI Retrieval Program is used to search the NDI file to determine whether a particular NDI death record qualifies as a possible record match with a particular user record. To qualify as a possible record match, both records must satisfy **at least one** of the following seven conditions or matching criteria, and the specified data items must agree on both records:

1. Social Security number
   (NOTE: For a search to be performed on the nine digits of a Social Security number, the NDI user's record must also contain EITHER a first and last name OR date of birth and sex; see preceding edit/reject criteria.)
2. Exact month and +/- 1 year of birth, first and last name
3. Exact month and +/- 1 year of birth, first and middle initials, last name
4. Exact month and day of birth, first and last name
5. Exact month and day of birth, first and middle initials, last name
6. Exact month and year of birth, first name, father's surname
7. **If the subject is female:** Exact month and year of birth, first name, last name (on user's record), and father's surname (on NDI record)

FIRST NAME or FIRST and MIDDLE INITIALS must match exactly to satisfy the above matching criteria. In the case of LAST NAME and FATHER'S SURNAME, however, the above matching criteria will be satisfied if there is agreement on names based on either their exact spelling or their New York State Identification and Intelligence System (NYSIIS) codes. The NYSIIS name coding procedure is intended to accommodate common misspellings of names by assigning a specified alpha character(s) to each character (or combination of selected characters) in a name; see examples in Exhibit 11.

With the liberal matching criteria described above, a given user record may generate possible record matches with more than one NDI record, especially when the subject's name is common (e.g., John Smith or Mary

Jones). The matching criteria are intended to maximize the number of **true** matches that can be found, especially for those NDI users unable to provide Social Security numbers. However, as a byproduct of these liberal criteria, a significant number of **false** matches (false positives) should be expected. The success of the NDI matching process will be determined by the following four factors:

- The effectiveness of the seven matching criteria.
- The quality and completeness of the data on your study subjects.
- The quality and completeness of the death certificate data in the NDI file.
- Your ability to assess the quality of the resulting matches.

NOTE: All NDI clients should submit as much of the information that is acceptable (see list on page 4) for their subjects. The objective is not to maximize the number of matches identified, but to maximize the number of "true matches" identified.

Responsibilities of NDI User

It is your responsibility to review the NDI output files, to assess the quality of each possible NDI record match listed, and to determine which possible matches are true matches. NDI records involved in matches, based on any of the seven matching criteria listed above, should only be considered **possible** matches. The format and content of the NDI output described in Chapter 3 are designed to assist you in determining which matches are **true** matches, which ones are **false** matches, and which are **questionable**, thus requiring further investigation.

If you requested an **NDI Plus** search (to obtain cause-of-death information), you will receive 10 files of NDI output, compared with eight files of output provided for a "routine" NDI search. **You are encouraged to familiarize yourself with each of these files as described in Chapter 3, especially the NDI Retrieval Report, before beginning to view or process them.** Call the NDI staff (301–458–4444) if you require assistance in processing the files or interpreting the output.

Options for Assessing NDI Output

The suggestions listed below should provide a few ideas on how to handle the wealth of information provided and how to assess the quality of matches that occur. As feedback is received from NDI users, this list of suggestions will be expanded in future revisions of this publication.

NOTE: To determine your **true** matches, refer only to the **REPORT** or **COMBINED** files. **As noted in Chapter 3, do not refer to the CAUSE or PRTCAUSE files until you have determined the true matches.** The **CAUSE** and **PRTCAUSE** files contain only the cause-of-death codes for the closer matches; however, many of these matches may also be false matches.

If you request only a routine NDI search, you will receive only the state and date of death and the death certificate numbers for each possible match. You will not receive the **CAUSE** or **PRTCAUSE** files with your NDI results.

Begin by developing specific criteria or matching conditions that must be satisfied for a match to be deemed a **true** match. Do the same to identify and eliminate **false** matches. Strive to minimize the number of matches that appear to be **questionable**. Some users may prefer to purchase death certificates for the questionable matches before making a final determination, but most users carefully assess the NDI results to decide which matches are the true matches and do not purchase death certificates.

Be prepared to receive a relatively large number of possible NDI record matches, many of which will be **false** matches.

If you submit a large volume of records for an NDI search, be prepared to develop a computer program to isolate **true** matches, to eliminate most **false** matches, and to minimize the number of **questionable** matches. To assist you in developing such a computer program, the NDI output contains the **COMBINED** file, entitled, "Combined File of Matching User and NDI Records." Refer to Chapter 3 for file format and coding information.

Records that match on a Social Security number will not always be true matches. A few of these matches may involve the record of a spouse. Checking further for agreements on sex, first name, or date of birth is helpful. Furthermore, on rare occasions a false match can occur when the Social Security number is reported or recorded wrong (on the user's or NDI record), causing a

match with another person's number that contains the same nine digits.

Matches generated solely by those criteria involving **date of birth** may be evaluated initially based on the extent to which the various digits of the **Social Security number** agree or disagree. The "NDI Retrieval Report" uses an "X" to show which digits of the Social Security number agree and a dash (—) to show which digits disagree.

Matches generated solely by those matching criteria involving **month** and **day of birth** may be initially evaluated based on the extent to which the years of birth disagree. The "NDI Retrieval Report" indicates the difference in years between the years of birth on an NDI record and a user record (e.g., +l year, 1 year, 15 years, and so on). In most instances, matches with greater than 10 years of death probably will be false matches. You will have to determine, however, your own specific criteria for eliminating such false matches based on agreements or disagreements on other data items.

The NDI Retrieval Program will generate and list **multiple** NDI record matches; that is, two or more NDI records may be matched with a given user record. This often occurs in the case of common names. The NDI records are ranked based on the number of NDI data items that agree with those on the user record. Keep in mind that none of the multiple matches is necessarily a **true** match; however, if the criteria you develop select one to be a true match, the remaining NDI record matches can be disregarded or eliminated. Furthermore, several users have commented that matches ranked fourth or greater seldom are true matches. Although this experience has not been analyzed and confirmed to date, you may want to consider either (a) eliminating all matches ranked fourth or greater or (b) applying looser elimination criteria for such matches.

The NDI Retrieval Program incorporates the use of phonetic codes for last names to enhance the marginal effectiveness of the NDI matching criteria. These phonetic codes are based on the New York State Identification and Intelligence System (NYSIIS). Accepting NYSIIS code agreements on names (rather than just exact agreements) tends to generate a few additional **true** matches; however, accepting NYSIIS code agreements on names can also result in the listing of numerous **false** matches. For example, the last name "**Johnson**" also translates to a NYSIIS code of "**Jansen.**" Similarly, the last names of "**Janzen**" and "**Jeanson**" also translate to the same NYSIIS code of "**Jansen**." (For more examples of last names associated with the same NYSIIS code, see Exhibit 11.) Consequently, NDI staff suggest that you apply stricter criteria when assessing matches resulting from NYSIIS agreements. Before determining that a **true** match has occurred, confirm whether sufficient agreements occur on other data items to warrant the designation of a true match.

If you request an NDI file search of only the most recent year of death available (or a few of the more current years of death), do not be surprised if you obtain several possible matches against NDI records having earlier years of death. These NDI records are **BELATED** records that were received from some states after the cutoff date for adding NDI records for a particular year of death. For example, a belated 2003 record will be flagged as being loaded with either the 2004 or 2005 deaths (depending on when it was received) and then will be stored in a special belated file which is **always** searched—regardless of the specific years you request to be searched. Consequently, if you submit records for repeat NDI searches as additional years of death are added to the NDI file, you will always have your records searched against any belated records that were missed in earlier searches. (Note that not all states submit belated records, and the average number of belated records added to the NDI file each year is less than 1,000.)

Probabilistic scores are assigned to each NDI record match. The scoring information appears in positions 172–178 of the **COMBINED** file. Refer to Appendix A for how the probabilistic scores are assigned. The assigned probabilistic scores and associated Status Codes (1 = true match and 0 = false match) are intended only as a guide for the user. The Status Code is also listed in the last column of the **REPORT** file. It is still the user's responsibility to determine which matches are true or false.

Once you have identified the true matches, the associated coded causes of death can be obtained from the **CAUSE** file or the **PRTCAUSE** file. The **CAUSE** and **PRTCAUSE** files contain ICD–9 and ICD–10 codes for the underlying cause of death, as well as the multiple causes of death (entity axis codes and record axis codes). Complete lists of these codes appear in the files containing the NDI Search Results. Explanations of the underlying cause-of-death and the multiple-cause codes are provided in Appendix B.

Be aware that the coded causes of death may not be available for some of the true matches that you selected. NDI is authorized to release only the coded causes of death for those matching NDI records that are **ranked first** in the list of possible matches. The only exception permitted is that coded causes can also be released

in those few instances when a matching NDI record is assigned a high-enough probabilistic score to be considered a **true** match (regardless of how the record is ranked). **(NOTE: If you selected *true* matches for which cause-of-death codes are not available on the CAUSE or PRTCAUSE files, contact NDI staff before attempting to purchase copies of death certificates. The staff may be able to provide you with the appropriate cause-of-death codes.)**

You may find it necessary to purchase copies of some death certificates whenever the causes-of-death are not available for some results.

Last Name Examples of NYSIIS Codes

LAST NAME = JOHNSON	LAST NAME = MORSE	LAST NAME = ANDERSON
NYSIIS = JANSAN*	NYSIIS = MARS*	NYSIIS = ANDARSAN*
JOHNSON	MORSE	ANDERSON
JAENSEN	MAERSCH	AANDERSEN
JAHNSEN	MARSAU	AANDERSON
JAHNSON	MARSAW	AMDERSON
JAHNSSEN	MARSCH	ANDAERSON
JAMSEN	MARSE	ANDARSEN
JAMSION	MARSH	ANODERSON
JANNSOHN	MARSHAE	ANDEERSON
JANSEN	MARSHAUS	ANDERSEN
JANSENIUS	MARSHAW	ANDERSENN
JANSHEN	MARSHIO	ANDERSENS
JANSON	MARZIO	ANDERSEON
JANSONIUS	MARZOA	ANDERSHIN
JANSSEN	MEARSE	ANDERSHONIS
JANSSON	MEHRZAI	ANDERSIN
JANSZEN	MERSI	ANDERSOHN
JANZAN	MEYERSHAW	ANDERSOM
JANZANO	MORSA	ANDERSONS
JANZEN	MORSCH	ANDERSOON
JEANSON	MORSEAU	ANDERSSEN
JEANSONNE	MORZE	ANDERSSOHN
JEHNSEN	MORZO	ANDERSSON
JEMSON	MOURSI	ANDERSUN
JENSEMA	MUERSCH	ANDERZEN
JENSEN	MURSAU	ANDERZHON
JENSSON	MURSCH	ANDERZON
JENZEN	MURSE	ANDORSON
JOHNSEN	MURSU	ANDIRSON
JOHNSION	MURZI	ANMDERSON
		ANNDERSON

* Not an exhaustive list of all the last names associated with the New York State Identification and Intelligence System (NYSIIS) code.

Probabilistic Scoring Approach for Assessing National Death Index Match Results

[**Caution**: National Death Index (NDI) users should be aware that submission records containing a significant number of missing data items (e.g., not collecting state of birth or state of residence) will have lower overall probabilistic scores and may, as a consequence, underestimate mortality for their cohort if additional clerical review of potential match records is not conducted.]

The probabilistic scoring technique described in this appendix is intended only to guide NDI users in determining which NDI record matches are likely to be **true** matches. Please read this appendix carefully before attempting to use the scores in your assessment of NDI matches. Note that the cutoff scores are fairly conservative, meaning that a **status code** of 1 implies a high probability that the NDI record is a **true** match (i.e., the study subject is assumed to be deceased). On the other hand, some portion of those NDI matches assigned a status code of 0 (assumed alive) may in fact also be true matches. The final responsibility of determining **true** and **false** matches rests with the NDI user.

NDI Matching Methodology

NDI is designed to facilitate health-related mortality studies. Researcher-supplied submission files are matched to the NDI computerized index of death record information compiled from death certificates submitted by state vital statistics offices to the National Center for Health Statistics (NCHS). Matching user submission records to NDI is a two-step process. In the first step, the NDI system selects potential death record matches based on a set of seven matching criteria. The second step consists of a scoring and classification procedure that results in the assignment of a probabilistic score and a suggested determination of final match status by NCHS. The probabilistic scoring applied to NDI matches is a modification of the approaches developed by Fellegi and Sunter (1969) and Rogot, Sorlie, and Johnson (1986); see Citations in this appendix.

Selecting NDI Records

NDI matches user submission records to death records based on seven criteria:

1. Social Security number (SSN)
   (NOTE: For a search to be performed on the nine digits of an SSN, the NDI user's record must at least contain EITHER a first and last name OR date of birth and sex.)
2. Exact month and +/- 1 year of birth, first and last name
3. Exact month and +/- 1 year of birth, first and middle initials, last name
4. Exact month and day of birth, first and last name
5. Exact month and day of birth, first and middle initials, last name
6. Exact month and year of birth, first name, father's surname
7. **If the subject is female:** Exact month and year of birth, first name, last name (**on user's record**) and father's surname (**on NDI record**)

Record matches between NDI records and user records are referred to as **possible matches**. An NDI record is selected as a possible match to a user record if it matches on **any one** of the seven criteria. *In the case of multiple NDI records returned for a given user record, the potential for a large number of false positives may occur. Of those matches listed, one may be a true match—but it is also possible that none may be a true match.* Alternatively, it is also possible that **no** NDI record will be selected for a given user record.

Indications of agreement between the user record and the NDI record are returned to the user for each possible match record. In addition to the data items involved in the seven matching criteria, NDI results return an indication of agreement for up to five additional data items:

- Age at death
- Race
- Marital status
- State of residence
- State of birth

Scoring and Classification of Potential Matches

Assessing the quality of possible matches and determining the best match for each user submission record requires a consistent approach. Each NDI possible match record is assigned a **probabilistic score**: the sum of the weights assigned to each of the identifying data items used in the NDI record match, where the weights reflect the degree of agreement between the information on the submission record and the NDI death record. NCHS developed the weights, known as binit weights, based on the frequency of occurrence of the identifying data items in NDI files for 1986–1991 and in the 1988–1991 National Health Interview Survey, a nationally representative survey of the noninstitutionalized U.S. population.

A **weight** is the base 2 logarithm of the inverse of the probability of occurrence of the characteristic based on the above files. For example, because males constitute 46.3% of the population aged 18 and over, the weight is $\log_2(1/.463) = 1.11$. Weights are constructed in a similar manner for race; last name; father's surname; birth month, day, and year; state of residence; and state of birth. The last-name weights have been modified for females. Because females have historically changed their surnames upon marriage, divorce, and remarriage, matching on surname only may produce a false nonmatch. The NDI returns an indication of either match or nonmatch on father's surname as well as last name. Because a person's father's surname does not change over time, this is used as auxiliary information for females. If last name does not match on the two records (the last-name weight is negative), the last-name weight is replaced with the father's surname weight if positive (matches); otherwise, the last-name weight is retained.

Because the distribution of middle initials differs by sex, sex-specific weights were constructed for middle initial. Weights for marital status were constructed to be jointly age- and sex-specific. Common first names, such as "John," which have a higher probability of occurrence, receive a lower binit weight than uncommon names, such as "Jonas." First-name weights are both sex- and birth-year-cohort specific because some secular trends occur in the assignment of first names. The weight assigned for SSN is a constant value of 30.

Weights are either positive or negative. If agreement exists between the user record and the NDI record for a particular identifying data item, the weight is positive. If there is no agreement, the weight is negative. Some items, such as year of birth, are assigned a weight of zero whenever the years are within a tolerance of +/-3 years. With the exception of middle initial, data items that are missing on the user's submission record, the NDI record, or both receive a weight of zero. A blank middle initial is considered a valid value and receives the appropriate weight. A New York State Identification and Intelligence System code (phonetic match) agreement is assigned only a partial weight because the exact spelling of the name did not agree. The score for each potential match is the sum of the weights for each individual data item:

$$\textbf{\textit{Score}} = W_{SSN} + W_{firstname\ x\ sex\ x\ birthyear} + W_{middleinitial\ x\ sex} + W_{lastname} + W_{race} + W_{sex} + W_{maritalstatus\ x\ sex\ x\ age} + W_{birthday} + W_{birthmonth} + W_{birthyear} + W_{stateofbirth} + W_{stateof\ residence}$$

After scoring the potential matches, each is categorized into one of five mutually exclusive classes. Whereas weighting and scoring consider the probability that the submission record and the NDI record share a particular value for the identifying items, the classes take into account which identifying items agree. Classes reflect the fact that some of the NDI identifying data items used in the matching criteria are more important for determining true matches than others. For example, as SSN is a key identifier in the matching process, each NDI record match is initially classified according to whether SSN is present and agrees (Class 1 or 2), is present but disagrees (Class 5), or is unknown (Class 3 or 4). Additionally, nonchanging identifying information is more important than information that can change over time. Many women, for example, assume their spouse's name at marriage, a common example of legitimate change over time. Birth surname, however, does not change and is thus an important matching variable for women. By contrast, state of residence and marital status may change over time and are, therefore, less important as classification variables.

The five classification groups developed by NDI are:

Class 1: Exact match on SSN (or at least eight digits), first name, middle initial, last name, sex, state of birth, birth month, and birth year.

Class 2: SSN matches on at least seven digits, and one or more of the other items from Class 1 may not match.
(NOTE: Some matched cases are moved from Class 2 to Class 5 because of an indication that the reported SSN belongs to the spouse. This includes those cases for which the SSN is known and matches, but the first name and sex do not agree.)

Class 3: SSN unknown but eight or more of first name, middle initial, last name, father's surname (for females), birth day, birth month, birth year, sex, race, marital status, or state of birth match.

Class 4: Same as Class 3 but fewer than eight items match.

Class 5: SSN is known but does not match.
(NOTE: Some matched cases are moved from Class 5 to Class 3 because of an indication that one of the SSNs (on the user record or on the death certificate) may have been reported incorrectly, but a significant number of other data items agree.

In this classification scheme, all of Class 1 matches are considered to be true matches, implying that the individuals are deceased. All of the Class 5 matches are considered false matches. Assignment of records falling into one of Classes 2, 3, or 4 as either true matches or false matches are based on score cutoff points within each class, as shown in the table. Records with scores greater than the cutoff scores are considered true matches, while records with scores lower than the cutoff scores are considered false matches. The recommended cutoff scores were determined on the basis of two calibration samples, with consideration given to jointly maximizing the **proportion** of records correctly classified and minimizing the **number** of records incorrectly classified.

NDI recommended and alternative cutoff scores are given in the table. The user is urged to adopt the recommended cutoff scores because they were chosen to provide overall optimal performance and are independent of any given study. If alternative cutoff scores are desired to conduct sensitivity analyses, they should be chosen within each class. The table provides estimates of the correct classification rates for each class under the assumption of 1,000 records within each class.

Table. Impact of Using Alternative Cutoff Scores

Class 2

	Deaths		Alive	
Score	N=926	Percent correct	N=74	Percent correct
34.5	924	99.8	7	9.1
39.5	924	99.8	9	12.1
44.5	**921**	**99.5**	**13**	**18.2**
49.5	917	99.0	18	24.2
54.5	906	97.8	20	27.3

Class 3

	Deaths		Alive	
Score	N=959	Percent correct	N=41	Percent correct
27.5	959	100.0	2	4.6
32.5	958	99.9	6	13.6
37.5	**946**	**98.6**	**23**	**56.1**
42.5	933	97.3	25	60.6
47.5	872	90.9	27	66.7

Class 4

	Deaths		Alive	
Score	N=281	Percent correct	N=719	Percent correct
22.5	191	67.8	622	86.6
27.5	173	61.6	684	95.2
32.5	**143**	**51.0**	**703**	**97.8**
37.5	119	42.4	712	99.0
42.5	70	24.9	718	99.8

NOTES: Sample size assumes 1,000 records in each class. N is the number of presumed correctly classified deceased and living people, based on a hypothetical sample of 1,000 people. The suggested cutoff score is the bolded middle score within each class. Use of an alternative cutoff score should be based on both the proportion and numbers of people correctly classified.

Example 1

This example of NDI weighting and scoring is based on a hypothetical person, with the characteristics as given in the following table:

Item	Value	Frequency	Weight
SSN	Unknown	NA	0.00
Last name	Robinson	0.00193	9.02
Middle initial	A	0.07748	3.69
First name	Leo	0.00140	9.48
Race	White	0.83509	0.26
Sex	Male	0.46329	1.11
Marital status	Married	0.05913	4.08
Birth day	10	0.03349	4.90
Birth month	October	0.08597	3.54
Birth year	1940	0.01418	6.14
State of birth	Florida	0.01552	6.01
State of residence	New York	0.06652	3.91

Because this record has an unknown SSN but at least eight or more of the following items are matched and have positive weights—first name, middle initial, last name, birth day, birth month, birth year, sex, race, marital status, or state of birth—this possible match record would be classified as a Class 3 match. The probabilistic match score then is the sum of the individual item weights, that is, 52.14. As 52.14 is greater than the Class 3 recommended cutoff score of 37.5, it would be assumed that this is a true match and the person is deceased.

Example 2

Using example 1, assume that birth day, marital status, and state of residence do not match but that the remaining items do match between the two records. This potentially matched record would be classified as a Class 4 match because less than eight of the following criteria are matched—first name, middle initial, last name, birth day, birth month, birth year, sex, race, marital status, or state of birth. In this example, the weights for birth day (4.90), marital status (4.08), and state of residence (3.91) would be negative, and the score would be 26.36. As this is less than the recommended cutoff score of 32.5 for Class 4, it would be assumed that this is a false match and that the person is not deceased.

Evaluation Studies

The NDI probabilistic scoring system was evaluated using two calibration samples. A calibration sample must have vital status information such as date and location of death, and ideally, death certificate number on the sample subjects based on sources independent of NDI. Two NCHS surveys met this criteria: National Health and Nutrition Examination Survey (NHANES) I Epidemiologic Follow-up Survey, also known as NHEFS, and the Longitudinal Study on Aging (LSOA).

The 14,407 people who participated in NHEFS (1971–1975) comprised the first calibration sample. Active follow-up was conducted on this sample to ascertain the vital status of the participants. Death certificates were obtained for people found to be deceased. NHANES is a large nationally representative survey and can be used as a calibration sample for developing a methodology for classification of potential NDI matches.

Because NDI was not begun until 1979, people who died prior to 1979 were eliminated from further consideration. Vital status was obtained independent of NDI by interviewer follow-back in 1982, 1986, and 1987. The NHEFS sample was then matched to NDI for 1979–1986. This yielded 5,393 records with potential matches to NDI and 6,672 records not involved in any matches.

The LSOA data set was used as a second calibration sample. LSOA was based on a subset of 1984 NHIS participants. The data used in this calibration sample are those participants aged 70 and over at the time of interview who were followed through August 1988. Vital status was obtained independent of NDI by interviewer follow-back in 1986 and 1988. Of the 7,541 people originally interviewed in 1984, 3,466 had potential matches with NDI (1984–1988), and 4,075 people were not involved in any match.

Subgroup Biases in Classification

The results of the evaluation study revealed biases in the classification of NDI match status for females and non-White people. The correct classification rate for females who were known to be deceased (94.0%) was about 2.5 percentage points poorer than for males (96.6%). This is due to linkage problems caused by changing surnames. Even though father's surname is being used to provide additional information, problems remain with correctly reporting and recording surnames in both the survey and on the death certificates. Both males and females had the same correct classification rates for living people.

Among non-White people, multiple problems, including lower reporting of SSNs and incorrect spelling or recording of ethnic names, can lead to underestimated mortality (or incorrectly classifying a true match as a false match). The correct classification rates for known decedents who are non-White dropped to 86% (89% in LSOA alone), while the classification rate for living people remained high at about 97%. The classification rate for female non-White people known to be deceased (84.7%) was about 3% lower than the classification rate for non-White male decedents (87.8%).

Differential reporting of SSN and correct name information results in a relatively large proportion of non-White potential matches classified as Class 4 matches. Class 4 consists of records with unknown SSNs and fewer than eight of the other items matching (due to errors or missing information). NDI users are urged to carefully evaluate the results of Class 4 matches, especially among matches for females and non-White people. Female and non-White matches assigned to Class 1, 2, 3, or 5 appear to have the same correct classification rates as those for white males.

Selected Citations

Fellegi IP, Sunter AB. A theory for record linkage. Journal of the American Statistical Association 64:1183–210. 1969.

Rogot E, Sorlie P, Johnson NJ. Probabilistic methods in matching census samples to the National Death Index. Chronic Dis 39:719–34. 1986.

For additional information on how NDI operates, refer to the NDI website for a bibliography titled, "Articles Describing the Performance of the National Death Index."

Coded Causes of Death

Data on causes of death are released by the National Center for Health Statistics (NCHS) in a variety of ways, including published reports, special tabulations to answer data requests, and public-use data sets. Since the inception of the multiple cause-of-death program in 1968, a public-use data set has been released for each data year. Each set contains a data record for all deaths processed by NCHS, including the underlying and multiple causes of death. For deaths occurring from 1979 forward, each record that NDI Plus links with the annual NCHS cause-of-death data set contains an underlying cause-of-death field, and two sets of multiple cause-of-death fields, which have been coded using the applicable *International Classification of Diseases* (ICD), published by the World Health Organization (WHO). Codes from the Ninth Revision (ICD–9) are used with data from 1979 through 1998. Beginning in 1999, codes from the 10th Revision (ICD–10) are used. The two sets of multiple cause-of-death fields can contain up 20 cause-of-death codes. These two sets of multiple cause-of-death fields are referred to as "entity-axis codes" and "record-axis codes," which are described later. The NDI Plus records contain the complete level of detail coded by NCHS, except where precluded by NCHS and state confidentiality restrictions, lack of data reliability, or when data are not available.

Lists of the ICD–9 and ICD–10 codes can be found in two separate files containing your NDI results. More complete, detailed ICD–9 and ICD–10 documents are available:

- ICD–9
- ICD–10

Underlying Cause-of-Death Classification

Since 1949, cause-of-death statistics have been based on the 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; International statistical classification of diseases and related health problems, 10th revision (ICD–10). 2008 ed. 2009.)

For each death, the underlying cause is selected from an array of conditions reported in the medical certification section on the death certificate. This section provides a format for entering the cause-of-death sequentially. The conditions are translated into medical codes through the use of the classification structure, and the selection and modification rules contained in ICD. Selection rules provide guidance for systematically identifying the underlying cause-of-death. Modification rules are intended to improve the usefulness of mortality statistics by giving preference to certain classification categories over others or to consolidate two or more conditions on the certificate into one classification category.

As a statistical datum, underlying cause-of-death is a simple, one-dimensional statistic; it is conceptually easy to understand and a well-accepted measure of mortality. It identifies the initiating cause-of-death and is, therefore, most useful to public health officials in developing measures to prevent the onset of the chain of events leading to death. The rules for selecting the underlying cause-of-death are included in ICD as a means of standardizing classification, which contributes toward comparability and uniformity in mortality medical statistics among countries.

Automated Selection of Underlying Cause of Death

Before data year 1968, mortality medical data were based on manual coding of an underlying cause-of-death for each certificate in accordance with WHO rules. Since 1968, NCHS has used automated coding systems to facilitate coding. Beginning with 2022 data, NCHS uses MedCoder, which combines new technologies with established technologies (e.g., Automated Classification of Medical Entities [ACME]).

The ACME system applies the same rules for selecting the underlying cause as would be applied manually by a nosologist; however, under this system, the computer consistently applies the same criteria, thus eliminating intercoder variation in this step of the process. The ACME computer program requires the coding of all conditions shown on the medical certification. These codes are matched automatically against decision tables that consistently select the underlying cause of death for each record according to the international rules. The decision tables provide the comprehensive relationships among the conditions classified by ICD when applying the rules of selection and modification. The decision tables were developed by NCHS staff on the basis of their experience in coding underlying causes of

death under the earlier manual coding system, and as a result of periodic independent validations. These tables periodically are updated to reflect additional new information on the relationship among medical conditions.

Multiple-Cause Data

The original scheme for coding conditions contained on the death certificate was designed with two objectives in mind. First, to facilitate etiological studies of the relationships among conditions, it was necessary to reflect accurately in coded form each condition and its location on the certification in the exact manner given by the certifier. Secondly, the codification needed to be carried out in a manner by which the underlying cause of death could be assigned through computer applications. The approach was to suspend the linkage provisions of ICD for the purpose of condition coding and to code each entity with minimum regard to other conditions present on the certification. This general approach is hereafter called **entity** coding.

Unfortunately, the set of multiple-cause codes produced by entity coding is not conducive to a third objective—the generation of person-based multiple-cause statistics. Person-based analysis requires that each condition be coded within the context of every other condition on the same certificate, and modified or linked to such conditions as provided by ICD–9. By definition, the entity data cannot meet this requirement because the linkage provisions distort the character and placement of the information originally recorded by the certifier.

Because the two objectives are incompatible, the Division of Vital Statistics creates from the original set of entity codes a new code set called **record-axis** multiple-cause data. Essentially, the axis of classification is converted from an entity basis to a record (or person) basis. The record-axis codes are assigned in terms of the set of codes that best describe the overall medical certification portion of the death certificate.

This translation is accomplished by a computer system called TRANSAX (translation of axis) through selective use of traditional linkage and modification rules for mortality coding. Underlying cause linkages which simply prefer one code over another for purposes of underlying cause selection are not included. Each entity code on the record is examined and modified or deleted as necessary to create a set of codes that are free of contradictions and the most precise within the constraints of ICD–9 and medical information on the record. Repetitive codes are deleted. The process may (1) combine two entity-axis categories into a new category, thereby eliminating a contradiction or standardizing the data, or (2) eliminate one category in favor of another to promote specificity of the data or resolve contradictions. The following examples from ICD–10 illustrate the effect of this translation:

Case 1: When reported on the same record as separate entities, cirrhosis of liver and alcoholism are coded to K74.6 (Other and unspecified cirrhosis of liver) and F10.2 (Mental and behavioral disorders due to use of alcohol; dependence syndrome), respectively. Tabulation of records with K74.6 would imply that such records had no mention of alcohol. A preferable code would be K70.3 (Alcoholic cirrhosis of liver) instead of both K74.6 and F10.2.

Case 2: If "gastric ulcer" and "bleeding gastric ulcer" are reported on a record, they are coded to K25.9 (Gastric ulcer, unspecified as acute or chronic, without mention of hemorrhage or perforation) and K25.4 (Gastric ulcer, chronic or unspecified with hemorrhage), respectively. A more concise code is K25.4, which shows both the gastric ulcer and the bleeding.

Entity-axis codes

Descriptions of the format and application of entity-axis data follow.

Entity-axis code format

Each entity-axis code is displayed as an overall seven-byte code with the following subcomponents:

- Line indicator: The first byte represents the line of the certificate on which the code appears. Six lines (1–6) are allowable, with the fifth denoting one or two written as "due to" beyond the four lines provided in Part I of the U.S. standard certificate of death. Line 6 represents Part II of the certificate.
- Position indicator: The next byte indicates the position of the code on the line [i.e., first (1), second (2), third (3), and up to the eighth (8) code on the line].
- Cause category: The next four bytes represent the ICD–9 or ICD–10 cause code.
- Nature-of-injury flag: ICD–9 uses the same series of numbers (800–999) to indicate nature-of-injury (N codes) and external-cause codes (E codes). This

flag distinguishes between the two with a one (1) representing nature-of-injury codes and a zero (0) representing all other cause codes. (NOTE: ICD–10 nature-of-injury codes are S000–T983 and do not have a flag of one (1) at the seventh position; the seventh position is blank.)

A maximum of 20 of these seven-byte codes are captured on a record for multiple-cause purposes. This may consist of a maximum of 8 codes on any given line, with up to 20 codes distributed across four or more lines depending on where the subject conditions are located on the certificate. Codes may be omitted from one or more lines (e.g., line 1 with one or more codes, line 2 with no codes, or line 3 with one or more codes).

In writing these codes, they are ordered as follows:

- line 1 first code, line 1 second code, line 1 third code, etc.
- line 2 first code, line 2 second code, line 2 third code, etc.
- line 3 first code, line 3 second code, line 3 third code, etc.
- line 4 first code, line 4 second code, line 4 third code, etc.
- line 5 first code, line 5 second code, line 5 third code, etc.
- line 6 first code, line 6 second code, line 6 third code, etc.

Fields without codes are left blank. The specifics of locations are contained in the record layout given in Chapter 3 of this User's Guide.

Entity-axis applications

The entity-axis multiple-cause data set is appropriate for analyses that require each condition to be coded as a stand-alone entity without linkage to other conditions or that require information on the placement of such conditions in the certificate. Within this framework, the entity data are appropriate to the examination of etiological relationships among conditions, accuracy of certification reporting, and the validity of traditional assumptions in underlying cause selection. Additionally, the entity data provide in certain categories a more detailed code assignment, which is linked in the creation of record-axis data. Where such detail is needed for a study, the user should selectively employ entity data. Finally, the researcher may not wish to be bound by the assumptions used in the axis translation process, preferring rather to investigate his or her own hypotheses.

By definition, the main limitation of entity-axis data is that an entity code does not necessarily reflect the best code for a condition when considered within the context of the medical certification as a whole. As a result, certain entity codes can be misleading or even contradict other codes in the record. For example, ICD–10 code K80.2 is "Calculus of gallbladder without cholecystitis." Within the framework of entity codes, this is interpreted to mean that Cholecystitis was not mentioned in the codable entity rather than that cholecystitis was not mentioned anywhere on the record. Consequently, tabulation of records with K80.2 would not accurately reflect the number of decedents having Calculus of gallbladder without cholecystitis. This illustrates the fact that under entity coding, the ICD–9 or ICD–10 titles cannot be taken literally. The user must study the rules for entity coding as they relate to his or her research prior to use of entity data. The user is further cautioned that the inclusion notes in ICD–9 or ICD–10 which relate to modifying and combining categories are seldom applicable to entity coding.

In tabulating the entity-axis data, codes may be counted with the resultant tabulation of an individual code representing the number of times the disease(s) indicated by the code appears in the data set. In this kind of tabulation of morbid condition prevalence, the counts among categories may be added together to produce counts for groups of codes. Alternatively, subject to the limitations stated above, people may be counted having mention of the disease represented by a code or codes. In this instance, it is not correct to add counts for individual codes to create person counts for groups of codes. Because more than one code in the researcher's interest may appear together on the certificate, totaling must account for higher-order interactions among codes. Up to 20 codes may be assigned on a record; therefore, a 20-way interaction is theoretically possible. All totaling must be based on mention of one or more of the categories under investigation.

Record-axis codes

The record-axis multiple-cause data are the basis for NCHS core multiple-cause tabulations. Descriptions of the **format** and **application** of record-axis data follow.

Record-axis code format

Each record (or person) axis code is displayed in five bytes. Location information is not relevant. The code consists of the following components:

- Cause category: The first four bytes represent the ICD–9 or ICD–10 cause code.
- Nature-of-injury flag: For ICD–9 codes, the last byte contains a 0 or 1, with the 1 indicating that the cause is a nature of injury category. (NOTE: ICD–10 nature-of-injury codes are S000–T983 and do not have a flag in the last byte; the last byte is blank. through 2017. Beginning with 2018 data, some maternal death records have a flag in the last byte of the second record axis code. In this case, the flag indicates that a response to a checkbox was the only indication that a pregnancy was involved. The code with the flag would have been the underlying cause without the checkbox.)

Again, a maximum of 20 codes are captured on a record for multiple-cause purposes. The codes are written in a 100-byte field in ascending code number (five bytes) order after the underlying cause (since 2005), with any unused bytes left blank. The record-axis codes are edited for rare causes and age/cause and sex/cause compatibility. Similarly, individual code validity is checked. The valid code set for record-axis coding is the same as that for entity coding.

Record-axis applications

The record-axis multiple-cause data set is the basis for NCHS multiple-cause tabulations. Location of codes is not relevant to this data, and conditions have been linked into the most meaningful ICD categories for the certification. The most immediate consequence for the user is that the record-axis codes already represent mention of a disease assignable to a particular ICD–9 or ICD–10 category through linkages. Secondly, the linkage implies that within the constraints of ICD, the most meaningful code has been assigned. The translation process creates a data set for the user that is edited for contradictions, duplicate codes, and imprecisions. (These contradictions, duplicate codes, and imprecisions exist in the entity-axis codes.) In contrast to entity-axis data, record-axis data also are classified in a manner comparable with underlying cause-of-death classification, which facilitates joint analysis of these variables. Similarly, they are comparable to general morbidity coding, where the linkage provisions of ICD are usually used. A potential disadvantage of record-axis data is that some detail is sacrificed in a number of the linkages.

The user can view the record-axis codes as literally representing the information conveyed in ICD category titles. While knowledge of the rules for combining and linking and coding conditions is useful, it is not a prerequisite to meaningful analysis of the data, as long as one is willing to accept the assumptions of the axis translation process. The user is cautioned, however, that due to special rules in mortality coding, not all linkage notes in ICD are utilized.

The user should proceed with caution in using record-axis data to count conditions, as opposed to people with conditions (the basis of NCHS multiple-cause tabulations). Linkages have been invoked and contradictions, duplicate codes, and imprecisions have been eliminated from the record-axis codes. As with entity data, person-based tabulations that combine individual cause categories must consider the possible interaction of up to 20 codes on a single certificate.

Updates to classification

Updates are extraordinary in ICD–9. For data year 1988, the decision tables mentioned previously were amended to incorporate minor changes to the classification for HIV infection (ICD–9 codes *042–*044) that originally had been implemented with data year 1987. Updates are more common in ICD–10. Additions and deletions of codes can be found at the end of Instruction manual part 9. Beginning in 2022, further changes were made in the record axis for HIV deaths involving other infections to retain detail on the other infections.

Military deaths

NDI also includes military deaths reported separately from deaths reported using death certificates. Information is originally reported on military forms that have three items related to manner, circumstances, and cause rather than the medical certification section of a death certificate. An ICD code for the underlying cause has been determined by NCHS for each of the specific values provided for the three items. This ICD code is provided for cause for these decedents.

ICD–9 and ICD–10 Recodes

For purposes of publishing data and compiling tables on causes of death from 1979 to the present, NCHS groups the data for selected ICD cause-of-death codes. Consequently, each cause-of-death code is provided with a "recode" to indicate with which group or data category a particular cause-of-death is associated.

Separate lists of recodes for ICD–9 and ICD–10 are provided.

Listed below are three sets of **ICD–9 recodes**. These recodes are included in your NDI results in data fields next to the ICD–9 underlying cause-of-death code:

- **282 recodes** of selected causes of death
- **72 recodes** of selected causes of death
- **61 recodes** of selected causes of **infant** deaths

IMPORTANT: ICD–10 recodes are used for deaths occurring **after** 1998. The ICD–10 recodes are positioned in the same data fields used for the ICD–9 recodes; however, the ICD–9 recodes change as follows:

- **282** (ICD–9) recodes become the **358** (ICD–10) recodes.
- **72** (ICD–9) recodes become the **113** (ICD–10) recodes.
- **61** (ICD–9) recodes become the **130** (ICD–10) recodes.

Copyright information

All material appearing in this report is in the public domain and may be reproduced or copied without permission; citation as to source, however, is appreciated.

Suggested citation

National Center for Health Statistics. National Death Index user's guide. Hyattsville, MD. 2024.

CS347384