

米国 NCHS 及びハーバード大学における NDI に関する現地調査

研究代表者 安村 誠司 福島県立医科大学医学部公衆衛生学講座教授

研究要旨 改正統計法は、十分な秘密保持を図ったうえで、統計データの利用を促進することとしている。日本の疫学研究・臨床研究のレベルを改善するためには、政府統計の利用申請課題では人口動態統計（死亡）の目的外利用を促進する方策を確立することは喫緊の課題である。死亡データベースのようなシステムのない日本における人口動態統計の現時点での効率的な研究活用の方策・課題を検討し、死亡者データベースの作成の必要性に関する提言を行う基礎資料を得ることが本研究の目的である。

そこで、欧米における死亡欧米各国の死亡者データベース National Death Index (NDI) の現状に関する情報収集の一環として、米国国立保健統計センターを訪問・視察し、NDI に関して、システムの概要、データベースの構成と運用体制、個人情報保護の措置など情報収集を行った。NDI 導入が有意義であることが確認された。また、ハーバード大学公衆衛生大学院を訪問し、新たな保健医療システムを行政システムとして導入する際の課題等について今後の方向性を討議した。

A. 研究目的

疫学・臨床研究を適切に行うには、対象者の生命予後を簡便かつ正確に把握する体制が不可欠であり、欧米では政府が死亡者データベースを作成して研究者に情報提供を行っている国が多数ある。

改正統計法は、十分な秘密保持を図ったうえで、統計データの利用を促進することとしている。日本の疫学研究・臨床研究のレベルを改善するためには、政府統計の利用申請課題では人口動態統計（死亡）の目的外利用を促進する方策を確立することは喫緊の課題である。

死亡データベースのようなシステムのない日本における人口動態統計の現時点での効率的な研究活用の方策・課題を検討し、死亡者データベースの作成の必要性に関する提言を行う基礎資料を得ることが本研究の目的である。

米国では、疾病管理予防センター（CDC; Center for Disease Control and Prevention）にある国立保健統計センター（NCHS; National Center for Health Statistics）が、死亡データベースである National Death Index (NDI) を作成・管理している。コホート研究、臨床研究（トライアルを含む）、疾病登録などで、研究者が対象者リストを NCHS に提出すると、個人の生死に関する情報が有料で提供される。これにより、簡便かつ正確に研究対象者全員の生命予後が把握できるようになった。NDI は米国の疫学研究や臨床研究のレベル・アップに大いに貢献している。しかし、その詳細についての情報は知られておらず、日本における NDI の導

入の可能性を検討する上で情報収集は必須と考えられ、現地調査を実施することとした。また、NDI という新たな保健医療システムを行政システムとして導入する際の課題等について意見聴取することを目的とした。

B. 研究方法

1. 3月3日に、米国 CDC の NCHS を訪問・視察し、NDI の Robert Bilgrad 部長から NDI に関する聞き取り調査を行った。なお、Mortality Statistics Branch の Robert N. Anderson 主任も同席し、情報集を行った。
2. 3月4日に、ハーバード大学公衆衛生大学院を訪問し、Michael R. Reich 教授から NDI の日本における導入する際の課題等について聞き取りを行った。さらに、Marcia Caldas de Castro 准教授からは、ブラジルにおける人口統計に関する情報収集を行った。

C. 研究結果

1. NCHS における NDI に関する情報収集
① NDI はもともと、米国で国レベルでの死亡データベースの必要性を議論する目的で 1976 年に NCHS に専門委員会が設置されたことが、その発端である。その後、1987 年 8 月 3 日に、“Final Report of the AD HOC WORKING TO DEVELOP PLANS AND PROCEDURES FOR THE POSSIBLE IMPLIMENTATION OF A NATIONAL DEATH INDEX” という形で、暫定委員会の最終報告が、NCHS に提出され、その中で、NDI

導入の必要性、重要性が指摘された。このような経過の後、1980年から部分的に、1982年からNDIを利用できる体制が構築された。使用できるデータは、1979年以降の死亡に関してである。

②NDIの組織・管理体制は、NCHSの建物の一角にあり、部屋はスタッフの部屋(3部屋)以外に、資料庫の部屋が共有であるのみである。スタッフは、Robert Bilgrad部長の他、アシスタント(助手)1名と、コンピュータプログラマー(非常勤)が1名の計3名のみであった。

③NDIの予算は、全体で、約\$2,200,000であるが、そのほとんどは、利用者からの申請費用(通常、基本料金\$350.00+1ケース・1年分で\$0.15×ケース数×年数で計算)で賄われており、補助金等が投入されている訳ではなかった。また、データは各州の統計局が収集したものをNCHSに集められ作成されており、データを集めた各州に、その収集した数に応じて、費用が支払われていた。

④申請者から申請書類が提出されると、National Death Index Advisory Panel、いわゆる、倫理委員会が申請書類を審査する。これに約3、4週間を要し、承認されると、申請されたデータに基づき、データの照合が行われ、データ量が少ない場合でも2、3週間は要するが、結果が報告される。概ね2カ月程度で結果が申請者に返されることになる。

⑤NDIへの申請はここ数年では、年間で、新規申請が約80~90件程度であり、継続申請が約200~250件程度ある。作業としては、コンピュータ上での作業であり、一定のロジックで行われるため、作業自体は大変ではないとのことであった。

⑥NDIでは大量のデータを扱う研究者の場合、申請費用がばくだいになる経済的負担を軽減する目的で、"VOLUME DISCOUNTS"と称した、値引きを実施している(資料2参照)

⑦現在、National Death Index User's Guideについては、改定の最中であり、Robert Bilgrad部長のご厚意で、入手することができた(資料3参照)。新たなUser's Guideでは、どのようにNDIを利用するかステップについて、詳細に、丁寧に記載されており、初心者でも十分利用できるように作られていることがわかる。

⑧個人情報の保護に関しては、申請者には、厳しい倫理的配慮が申請時点でサイン(署名)を求められており、これを順守しなければ、

罰せられる仕組みになっていた。

2. ハーバード大学における情報収集

①Michael R. Reich教授からの情報収集の中で、もっとも重要と考えられた点の一つが、なぜ、米国ですでに実施され有用であることが日本でも研究者、特に、疫学研究者ではよく知っているのにもかかわらず、導入が日本では進まないのかという疑問であった。新たな制度・システムを導入することに対する消極的な態度は、日本の官僚組織の歴史的な特徴ともいえるものであり、容易に変更することは困難であるとの考え方に立つ必要があることが指摘された。

②日本の研究者、特に、疫学研究者の中ではNDIが知られているとしても、他の臨床研究者や、医師を含めた保健医療関係者にその意義や有用性が認知されているかは疑問であるとの意見が出された。

③日本においては、死亡情報が含まれる人口動態統計については厚生労働省の所管であるが、これら政府統計の目的外使用の申請先は、総務省であり、管轄が異なっている。このような、政府統計に関する管轄の構造的な問題も目的外使用の煩雑さの原因の一つであろう。また、厚生統計の活用意義についての理解が、総務省に十分あるかどうかという点も、目的外使用の許可に時間を要する原因とも考えられる。

④ブラジルの死亡データベースに関しては、NDIと同じようなシステムはないと考えられた。しかし、少なくとも、匿名化された死亡データベースはブラジル政府の厚生省のホームページにアドレスが掲載されており、それにアクセスすることで、データを入手することは可能であることがわかった。

D. 考察

1. NCHSにおけるNDIに関する情報収集

①米国においても、まず初めは、他分野の多数の構成員からなる暫定的な委員会が結成され、その報告として、死亡データベースの必要性が提言として発表されたことがきっかけとなっていた。日本においては、本研究班がまさに同様の役割を持つ委員会としての機能を果たすことが重要であることが改めて認識できた。

②NDIの予算では、多岐にわたり、その状況を把握したが、NDIがスタートする際には、当然、コンピュータプログラムの作成を含め初期投資はある程度必要であることは確かであろうが、ランニングコストを含め、利用者

からの申請費用で賄われており、費用的には問題はそれほど問題にはならないのではないかと考えられた。

③NDI の組織、スタッフ、運用体制などでは、場所は広い場所は必要なく、機材等としては、コンピュータがあれば、事足りることがわかった。また、申請書類を保管する書庫がいくつかあれば、数十年分を保管するには十分であり、スペースを取らないこともわかった。また、人員も NDI で3名のみであり、多人数を確保する必要もないことから、多大な負担のもとでの開始・運営にはならないと確信できた。

④個人情報保護については、その順守が申請者には、厳しく求められており、個人情報保護の点からは問題は、日本においても NDI を導入する際には、このような厳しい対応が必要であると考えられた。

2. ハーバード大学における情報収集

①日本で所管する厚生労働省が所管する人口動態統計の目的外使用を認可するのが総務省であるという問題は、ここで議論しても解決方法がなく、むしろ、NDI を導入の意義・有用性を厚生労働省に理解してもらうことが重要である。

②また、疫学研究者は研究者全体の中ではその数が少なく、また、政治的な力もほとんど持ちえないの実態である。疫学研究者のみが NDI の意義や有用性を訴えても、その力は弱く、他の学協会の理解と協力を得ることが必須と考えられた。具体的には、「医学の進歩、発展、つまりは、国民の医療・福祉の向上に資する」という点で、日本医学会の他、日本医師会等からの協力も重要であろう。また、マスコミや、企業、各種経済団体、労働組合などにとっても NDI 導入の恩恵がある点を周知する必要がある。むしろ、国民の支持が大前提であることは言うまでもなく、国民への情報提供、さらには、国民からの後押しも忘れてはならない。

③隣国、韓国ではすでに NDI が導入されている。日本の方が保健医療システムが進んでいると、政府、厚生労働省は考えているかもしれないが、実態は決してそうではない。早急に導入することで、保健医療水準の向上に資する対策を立てる必要がある。

E. 結論

1. NCHS における NDI に関する情報収集から、NDI は研究の促進の点からその意義がきわめて高く、NDI の導入の恩恵は研究者が得るのではなく、国民全体が得られるということがわかった。NDI の導入に関して初期投資は当然日つであるが、そのシステムの維持についての、人・物・金の点は大きな問題とはならないと考えられた。
2. ハーバード大学における情報収集から、NDI 導入における問題を解決するためには、直接の管轄である厚生労働省担当者、NDI の意義、必要性を十分に理解してもらうことがまず重要である。また、NDI の導入の必要性については、国民全体の保健医療水準の向上に資するものであるとの視点から、日本医師会、マスコミ、各種経済団体等、国民各界からの支持が得ら得るような、活動を行う必要がある。

F. 健康危険情報

なし

G. 研究発表

1. 論文発表
なし
2. 学会発表
なし

H. 知的財産権の出願・登録状況 (予定を含む)

1. 特許取得
なし
2. 実用新案登録
なし
3. その他
なし

研究協力者

中村好一 自治医科大学医学部公衆衛生学講座

顧問

児玉和紀 主席研究員 (放射線影響研究所)・
日本疫学会理事長



NATIONAL DEATH INDEX PLUS VOLUME DISCOUNTS



For searches of subjects with UNKNOWN vital status:

Number of Records	Current Fees *	Discounted Fees	Year Caps **
First 100,000	\$0.21	\$0.21	No Caps
Next 2,400,000	\$0.21	\$0.21	2
Each additional record	\$0.21	\$0.025	2

* Current fees are \$0.21 per record per year of death searched.

** After the first 100,000 records, discounted fees have a 2-year cap for "initial" NDI searches, regardless of the number of years of death the subjects' records need to be searched.

The 2-year caps are NOT offered for "repeat" searches of the same records for subsequent years of death. It is presumed that repeat searches would be conducted annually (or at least every 2 years) to keep a user's database current.

For searches of KNOWN decedents:

Number of Records	Current Fees	Discounted Fees	Year Caps
First 100,000	\$5.00	\$5.00	1
Next 400,000	\$5.00	\$1.00	1
Each additional record	\$5.00	\$0.05	1

	Examples of Discounted Fees by Record Counts *		
	500,000	1,000,000	5,000,000
* First 100,000	\$105,000	\$105,000	\$105,000
Next 2,400,000	\$168,000	\$378,000	\$1,008,000
All additional records	--	--	\$125,000
Total cost:	\$273,000	\$483,000	\$1,238,000

* The above examples assume all records would have needed a 5-year NDI search, consequently the fees for the first 100,000 records are 100,000 x \$0.21 x 5 years. The calculation of fees for all additional records have a 2-year cap.

Total cost: \$900,000 \$925,000 \$1,125,000 \$1,375,000

	Examples of Discounted Fees by Record Counts		
	500,000	1,000,000	5,000,000
First 100,000	\$500,000	\$500,000	\$500,000
Next 400,000	\$400,000	\$400,000	\$400,000
All additional records	--	\$25,000	\$225,000
Total cost:	\$900,000	\$925,000	\$1,125,000

Total cost: \$900,000 \$925,000 \$1,125,000 \$1,375,000

Qualifying for volume discounts:

1. The organization must submit an NDI Application Form and obtain approval for the project.
2. EACH submission must include at least 100,000 records of UNKNOWN or KNOWN.
3. No record level information obtained from the NDI shall be re-released in an identifiable form to any outside organization.

NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI
NDI

DRAFT: February 23, 2009

National Death Index

User's Guide

NOTE: Effective 2/23/2009, this draft document replaces the following two NDI documents:

- *NDI User's Manual*
- *NDI Plus: Coded Causes of Death*

Before we finalize this document, we would appreciate your input. Please provide comments, questions, and suggestions to:

Robert Bilgrad
301-458-4101
RBilgrad@cdc.gov

Department of Health and Human Services
Centers for Disease Control and Prevention
National Center for Health Statistics

National Death Index User's Guide

TABLE OF CONTENTS

Chapter 1	How to Use the NDI: Steps in the Process	4
Chapter 2	Preparing Your Records: Record Layout & Coding Specifications	5
	Exhibit 1 NDI User Records, Sample Format	6
	Exhibit 2 NDI User File Format and Coding Specifications	7
	Exhibit 2A Coding Instructions for Names	11
	Exhibit 2B State (or Foreign Country) of Residence, Birth and Death	13
Chapter 3	Your NDI Results	14
	General description of NDI output files	15
	Exhibit 3 NDI Output File Descriptions	16
	[EDITS] User File Edit Results	17
	Exhibit 4 Edits (example)	18
	[SUMMARY] Summary Retrieval Statistics	20
	Exhibit 5 Summary Retrieval Statistics (example)	21
	[REPORT] NDI Retrieval Report	23
	Exhibit 6 NDI Retrieval Report (example)	24
	Exhibit 6A Retrieval Report (Abbreviations and Symbols)	25
	[COMPRESS] Compressed NDI Retrieval Report	27
	Exhibit 7 Compressed NDI Retrieval Report (example)	28
	[COMBINED] Combined File of Matching User and NDI Records	29
	Exhibit 8 Record Layout for the Combined File	30
	[CAUSE] Cause of Death File	34
	Exhibit 9 Record Layout for the Cause File	37
	[PRTCAUSE] Cause of Death Report	45
	Exhibit 10 Cause of Death Report (example)	46
	[REQFORMS] Death Certificate Request Forms	47
	Exhibit 11 Death Certificate Request Form (example)	49
	Exhibit 12 U.S. Standard Certificate of Death	50
	Exhibit 12A Cause-of-Death – Background, Examples, & Common Problems	51
	[MATCH], [NOMATCH], [REJECTS]	52
	Matching User Records, Non-Matching User Records & Rejected User Records	

Chapter 4 Assessing the NDI Output	53
Review Your NDI Edit Results <i>First</i>	53
Matching Criteria	53
Responsibilities of the NDI User	54
Suggestions for Assessing the NDI Output	54
Exhibit 13 Last Name Examples of NYSIS Recodes	57
Appendix A A Probabilistic Scoring Approach for Assessing NDI Match Results	58
NDI Matching Methodology	58
Selecting NDI Records	58
Scoring and Classification of Possible Matches	59
Table 1 The Impact of Using Alternative Cutoff Scores	61
Example 1	62
Example 2	62
Evaluation Studies	62
Subgroup Biases and Classification	63
References	64

CHAPTER 1 - How to Use the National Death Index (NDI): Steps in the Process

1. Download the NDI Application Form from the NDI website. (*Allow approximately two to three months for your application to be reviewed and approved.*)
2. Call NDI staff for guidance (301-458-4444) before completing the application form.

Note: *To obtain cause of death codes for any deaths that may occur in Florida or New York City, you will also need to submit confidentiality forms required by Florida and New York City. Have NDI staff email you these forms. These forms must be signed and notarized and attached to your **final** NDI Application Form. Contact NDI staff if you decide to "waive" Florida and/or New York City.*

3. Email your **unsigned** NDI Application Form to ndi@cdc.gov. (This submission does not need to be accompanied by a copy of your IRB approval document or by your Florida and/or New York City forms.)
4. NDI staff will respond by emailing your assigned NDI number. Then, in about a week NDI staff will let you know whether your application might need revisions before you submit your **final** application.
5. Express mail your **final, signed** NDI Application Form and IRB approval to NCHS. (Include signed, notarized forms for Florida and New York City. NDI staff will forward these forms to Florida and New York City.) Email us at ndi@cdc.gov or call 301-458-4444 to confirm that we received your package.
6. Your application will be mailed to a 12-member panel for review and comment. (Your Florida and New York City forms will be forwarded to the appropriate vital statistics offices for approval along with a copy of your NDI Application Form.)
7. While your application is being reviewed, you can begin preparing your study subjects' records for submission (on a CD or diskette). Carefully follow the specifications in Chapter 2 of this [User's Guide](#). (Note: We recommend that you password protect your files using **WinZip** or any self-decrypting software.)
8. You will receive a letter of approval and a blank **NDI Transmittal Form** when your application is approved.
9. Express mail your file(s) of study subjects' records to NCHS. Include a separate **NDI Transmittal Form** for each file, a *Worksheet for Calculating NDI Charges*, and your check or purchase order. Then email us your password and your express mail tracking number.
10. NDI staff will express mail your NDI search results in about two weeks.
11. Assess your results to identify the **true** matches using your own assessment criteria and/or the suggested NDI assessment criteria. You also have the option of purchasing death certificates directly from the state vital statistics offices—for all **true** matches or just for those matches you feel are questionable.
12. If additional NDI searches are needed in the future for the **same** study or project, first mail NCHS a signed **NDI Repeat Request Form**. (You will receive a new *Repeat Request Form* each time you receive your NDI results.) As soon as you receive a letter approving your repeat request (usually within 2 weeks), you may submit your new files for an NDI search.

NOTE: Whenever you have questions, please call us on 301-458-4444 or email us at ndi@cdc.gov. When emailing, always include your phone number and assigned NDI number.

CHAPTER 2 - Preparing Your Records: Record Layout and Coding Specifications

1. Records of study subjects must be created in a **standard TEXT FILE** (flat file) format.
2. Each record must contain **exactly 100 positions**.
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**
4. **DO NOT USE ANY DELIMITERS SUCH AS TABS OR COMMAS TO SEPARATE THE DATA ELEMENTS.**
5. Each data element must be put 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. An **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.
10. Only submit your records on a CD-ROM or a standard 3.5 inch diskette. If you put more than one file on the CD or diskette, it is helpful to use your assigned NDI numbers in your file names. Also print your assigned NDI numbers on the outside of your CD or diskette. **We strongly recommend that you password protect your file(s) using WinZip or any self-decrypting software.**
11. Express mail your files(s), Transmittal Form, fee worksheet, and payment to:

NATIONAL DEATH INDEX
National Center for Health Statistics
3311 Toledo Road, Room 7318
Hyattsville, MD 20782
Phone: 301-458-4444
12. E-mail ndi@cdc.gov to confirm that your file(s) was received. Include your NDI number, the date the file was sent, and your express mail tracking number. If your file(s) is password protected, include your password(s) in the e-mail.

Exhibit 1: NDI User Records, Sample Format

Create your records in a **standard text file format** with **100 positions** per record and with no tabs, commas or other delimiters. We understand that your study subjects' records are not stored in text files. We recommend that you use software (such as SAS) that will enable you to create a text file and to position each data element in the exact record positions described below. If you are only submitting a small number of records and plan to key your data into a text file, we suggest you use text file software such as Microsoft **Notepad**. In *Notepad* you can click on VIEW and activate the STATUS BAR which enables you to see the line and column numbers. Change the font to a proportionately spaced font such as LUCINDA CONSOLE. This will insure that each data item lines up properly.

Positions of NDI data items

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

Sample data format

1	2	3	4	5	6	7	8	9	0
12345678901234567890	12345678901234567890	12345678901234567890	12345678901234567890	12345678901234567890	12345678901234567890	12345678901234567890	12345678901234567890	12345678901234567890	12345678901234567890
WASHINGTON	GEORGE	T285160856	04061731	WASHINGTON		075121013059A536a56	060183		
BORDON	Elizabeth	Q344654136	01211922	Adams		075121013059387	052880		
Franklin	Benjamin	D786782104	05221701	Franklin		10111216NH49596DUP	120998		
Munoz	Christine	262702171	11231926	ARIAS		G29211NVCA62345	122296		
ARBERTHA	Joyce	A344454644	07041961	Brown		035222MDMD59387	100889		

Exhibit 2: NDI User File Format and Coding Specification

DATA ELEMENTS	POSITION	CODE STRUCTURE	CODING INSTRUCTIONS	
1. Name of Person in the Study Group			* See EXHIBIT 2A for instructions for coding surnames, first names, and middle initials.	
Last Name	1-20	Alpha, left justified	* 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 Name	36	Alpha or blank		
2. Social Security Number	37-45	Alpha/Numeric or blank	<p>* Must have 9 digits. Do not drop any leading zeroes.</p> <p>* If less than 9 or partial number, leave the field blank.</p> <p>* If more than 9 digits, enter the first 9 digits. * If two Social Security numbers, enter the first or create a duplicate record.</p> <p>* If the number is unknown or not provided, leave the field blank.</p> <p>* If SOCIAL SECURITY NUMBER is not provided, the MONTH and YEAR OF BIRTH must be provided or the record will automatically be rejected.</p>	
3. Date of Birth			* The months of January-September and the days 1-9 must have leading zeros.	
Month	46-47	January	01	<p>* If either MONTH or YEAR of birth is not provided, the SOCIAL SECURITY NUMBER must be provided or the record will automatically be rejected.</p> <p>* If only a person's YEAR of birth is known (and MONTH of birth is unknown), you should consider the merits of preparing 12 duplicate records, one for each possible MONTH of birth. This is especially important when the SOCIAL SECURITY NUMBER is not known. (There is no charge for duplicate records.)</p>
		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 if unknown		

DATA ELEMENTS	POSITION	CODE STRUCTURE	CODING INSTRUCTIONS																		
4. Father's Surname	54-71	Alpha, left justified or blank if unknown	<p>* Follow the coding instructions for "surnames" as specified in EXHIBIT 2A.</p> <p>* For females the provision of FATHER'S SURNAME is encouraged.</p> <p>* For males FATHER'S SURNAME is preferred only when it differs from the LAST NAME.</p>																		
5. Age at Death																					
Unit	72	<table border="1"> <tr> <td>Year: less than 100</td> <td>0</td> </tr> <tr> <td>Years: 100 or more</td> <td>1</td> </tr> <tr> <td>Months</td> <td>2</td> </tr> <tr> <td>Weeks</td> <td>3</td> </tr> <tr> <td>Days</td> <td>4</td> </tr> <tr> <td>Hours</td> <td>5</td> </tr> <tr> <td>Minutes</td> <td>6</td> </tr> <tr> <td>Age unknown</td> <td>blank or 9</td> </tr> </table>	Year: less than 100	0	Years: 100 or more	1	Months	2	Weeks	3	Days	4	Hours	5	Minutes	6	Age unknown	blank or 9	<p>* 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.</p> <p>* The "unit" field identifies the type of units for which age at death is measured.</p> <p>Example: Age 65 Unit = 0 Number of Units = 65</p> <p>Example: Age 105 Unit = 1 Number of units = 05</p>		
Year: less than 100	0																				
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	<table border="1"> <tr> <td>Enter age</td> <td>01-99</td> </tr> <tr> <td>Age unknown</td> <td>blank or 99</td> </tr> </table>	Enter age	01-99	Age unknown	blank or 99	<p>Example: Age 8 months Months Unit = 2 Number of units = 08</p> <p>Example: Age unknown Unit = blank or 9 Number of units = blank or 99</p>														
Enter age	01-99																				
Age unknown	blank or 99																				
6. Sex																					
	75	<table border="1"> <tr> <td>Male</td> <td>1 or M</td> </tr> <tr> <td>Female</td> <td>2 or F</td> </tr> <tr> <td>Unknown</td> <td>blank or 9</td> </tr> </table>	Male	1 or M	Female	2 or F	Unknown	blank or 9													
Male	1 or M																				
Female	2 or F																				
Unknown	blank or 9																				
7. Race																					
	76	<table border="1"> <tr> <td>White</td> <td>1*</td> </tr> <tr> <td>Black</td> <td>2</td> </tr> <tr> <td>Indian</td> <td>3*</td> </tr> <tr> <td>Chinese</td> <td>4</td> </tr> <tr> <td>Japanese</td> <td>5</td> </tr> <tr> <td>Hawaiian (includes part Hawaiian)</td> <td>6*</td> </tr> <tr> <td>Other nonwhite</td> <td>7*</td> </tr> <tr> <td>Filipino</td> <td>8</td> </tr> <tr> <td>Other Asian or Pacific Islander</td> <td>0</td> </tr> </table>	White	1*	Black	2	Indian	3*	Chinese	4	Japanese	5	Hawaiian (includes part Hawaiian)	6*	Other nonwhite	7*	Filipino	8	Other Asian or Pacific Islander	0	<p>* <u>White</u> (includes Mexican, Puerto Rican and all Caucasian)</p> <p>* <u>Indian</u> (includes American, Alaskan, Canadian, or Mexican Indian, Eskimo, and Aleut)</p> <p>* <u>Hawaiian</u> (includes part Hawaiian)</p> <p>* <u>Other nonwhite</u> (includes Cajun and Creole)</p>
White	1*																				
Black	2																				
Indian	3*																				
Chinese	4																				
Japanese	5																				
Hawaiian (includes part Hawaiian)	6*																				
Other nonwhite	7*																				
Filipino	8																				
Other Asian or Pacific Islander	0																				

DATA ELEMENTS	POSITION	CODE STRUCTURE	CODING INSTRUCTIONS	
		Unknown, not stated, or not classifiable	blank or 9	
8. Marital Status	77	Never married/ single	1*	
		Married	2*	
		Widowed	3	
		Divorced	4	
		Unknown	blank or 9	
			* <u>Other Entries</u> Annulled 1 Separated 2 Common law marriage 2	
9. State of Residence	78-79	ALPHA and/or NUMERIC codes	* See EXHIBIT 2B for the ALPHA and/or NUMERIC codes which may be used. * 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.	
		States		01-51
		Puerto Rico		52
		Virgin Islands		53
		Guam		54
		Canada		55
		Cuba		56
		Mexico		57
		Remainder of World		59
		Unknown		blank or 99
10. State of Birth	80-81	(Same code structure as "State of Residence" above)	See EXHIBIT 2B for the ALPHA and/or NUMERIC codes which may be used.	

DATA ELEMENTS	POSITION	CODE STRUCTURE	CODING INSTRUCTIONS
11. Identification (ID), Control, or Sequence Number (OPTIONAL)	82-91	Any combination of ALPHA/NUMERIC; left justified. If no numbers are assigned, leave this field BLANK.	<p>* This is an OPTIONAL field to assist NDI users in identifying the records they submit. Insert identification numbers or control numbers assigned to persons in the study OR assign "sequence numbers" based on the order that records appear on the file to be submitted to NCHS.</p> <p>* 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).</p> <p>* 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., NCHS 1, NCHS 2, NCHS 3, etc.</p>
12. Optional User Data	92-97	Any combination of ALPHA/NUMERIC; left justified If this field is not used, leave it BLANK	<p>This OPTIONAL field may be used for any additional information on study subjects. Examples of possible uses are:</p> <p>* Date or year of last contact (e.g., 011584 or just 84 or just 1984)</p> <p>* For KNOWN decedents, date or year of death (e.g., 023091 or just 91, just 1991)</p> <p>* 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)</p> <p>* Coded information to identify special subgroups of study subjects</p> <p>* To flag DUPLICATE or ALIAS records</p> <p>The NDI computer output will contain whatever the user decides to include in this field.</p>
13. Blank Field	98-100		

EXHIBIT 2A: 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. The NDI edit program will remove the space or the hyphen.
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: SMITH-LEE
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, you should consider the merits of creating a duplicate or alias record(s), *one record for each surname*. 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. Suffixes or generational identifiers such as II, III, Jr., or Sr. will be removed by the NDI edit program.
7. For names such as Sister Mary Lawrence, enter LAWRENCE as the last name, SISTER MARY as the first name, and a BLANK for middle initial.

EXHIBIT 2A: CODING INSTRUCTIONS FOR NAMES (continued)

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 nick name, or (c) uses the shortened form of a presumed legal name (for example, Bob for Robert, Liz for Elizabeth, Ben for Benjamin) you should 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, 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.

EXHIBIT 2B: STATE (OR FOREIGN COUNTRY) OF RESIDENCE, BIRTH AND DEATH

State	Alpha Code	Numeric Code		State	Alpha Code	Numeric Code
Alabama	AL	01		New Jersey	NJ	31
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		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		Unknown		99
New Hampshire	NH	30				

CHAPTER 3 - Your NDI Results

This chapter describes the different files you will receive on the NDI Search Results CD. But before you actually begin working with these files, we strongly recommend that you also review Chapter 4 for suggestions on how to assess your results.

IMPORTANT REMINDER! 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 applicant (plus other parties involved in the study or project) provides NCHS with confidentiality assurances that information obtained from the National Death Index, from state death records, and from death record followback investigations will not be published or released in any form if a particular individual or establishment supplying the information or described in it is *identifiable*.

Furthermore, the applicant has assured NCHS that the *identifying* information:

- (1) will be used only for statistical purposes in medical and health research.
- (2) will not be used as a basis for legal, administrative or other actions which may directly affect those particular individuals or establishments as a result of their specific identification in the study or project.
- (3) will be used only for the study or project described in the approved NDI Application Form.

As an added precaution, it is strongly recommended that you destroy any identifiable death record information (and especially coded causes of death) associated with those NDI record matches which 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 the National Center for Health Statistics whenever your organization receives (or expects to receive) a subpoena or court order for identifiable death record information obtained as a result of using the NDI. NCHS will do everything possible to assist in protecting the confidentiality of the identifiable information.

General description of NDI output files

- The NDI system generates *nine* basic files of output for each NDI search. If you opted for our NDI *Plus* service, you will receive *two* additional files containing the cause of death codes for your better NDI matches. In addition, you will receive two files containing complete lists of the ICD -9 and ICD-10 cause of death codes.
- The eleven NDI output files are in a standard text file format and are stored on a CD -ROM entitled **NDI Search Results**. A brief description of each of the NDI output files is presented in **Exhibit 3**.
- The following six NDI output files are formatted to be **printed** at your discretion:

User File Edit Results (*EDITS*)
Summary Retrieval Statistics (*SUMMARY*)
NDI Retrieval Report (*REPORT*)
Compressed NDI Retrieval Report (*COMPRESS*)
Cause of Death Report (*PRTCAUSE*)
Death Certificate Request Forms (*REQFORMS*)

- After you review the content and size of each of these files, you can decide which ones you want to print either partially or completely. When printing these files, be sure to use the *landscape* format and, if necessary, reduce the size of your margins and/or fonts to capture all the information.
- If you decide that the NDI Retrieval Report (*REPORT*) is too large to print, you might decide to only print the Compressed Retrieval Report (*COMPRESS*). This file contains the same information as the *REPORT* file; however, the column headings and skipped lines are removed to make the printout more manageable.
- The following 2 files are created by the NDI 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 (*COMBINE*)
Cause of Death File (*CAUSE*)

- The file labeled *MATCH*, *NOMATCH* and *REJECTS* only contain the 100-character records you submitted. 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 package will also contain the CD or diskette you submitted for the NDI search. Your identifying information is backed up on CDC's computers and then deleted after 60 days.
- A description of each of the 11 NDI output files is presented in the remainder of this chapter.

EXHIBIT 3 – NDI output file descriptions

All file names contain the user's NDI Search Number preceded by the letter "Y" and followed by a file identifier (e.g., Ynnnnnn.EDITS). All these files are *text* files but may not have a ".txt" extension at the end of their file names.

FILE NAME	FILE DESCRIPTION	RECORD LENGTH
Ynnnnnn.EDITS	User File Edit Results (EDITS) – List of rejected user records and statistics on the completeness and quality of each data item.	132
Ynnnnnn.SUMMARY	Summary Retrieval Statistics (SUMMARY) – Statistics on the number of matches, multiple matches, and matches by State of death.	132
Ynnnnnn.REPORT	NDI Retrieval Report (REPORT) – User records and corresponding NDI record matches; in pre-printed format with column headings. Intended solely for the visual inspection of matches.	132
Ynnnnnn.COMPRESS	Compressed NDI Retrieval Report (COMPRESS) – Same as the <i>REPORT</i> file but with column headings and line spacing <i>eliminated</i> to reduce the amount of computer paper printed. Intended solely for the visual inspection of a large number of matches.	132
Ynnnnnn.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. Intended solely for designing computer programs to assess a large number of matches. Contains probabilistic scoring in positions 172-178.	178
Ynnnnnn.CAUSE *	Cause of Death File (CAUSE) – The file only contains those possible NDI record matches which are ranked first or those which are assigned a high enough probabilistic score. Positions 1-178 are identical to those in the <i>COMBINED</i> file. Positions 180-438 contain the coded causes of death.	438
Ynnnnnn.PRTCAUSE *	Cause of Death Report (PRTCAUSE) – This file is intended to be printed solely for the visual inspection of the coded causes of death for the same NDI record matches included in the <i>CAUSE</i> file.	132
Ynnnnnn.REQFORMS	Death Certificate Request Forms (REQFORMS) – List of NDI record matches by STATE to be used when requesting copies of death certificates. There is a separate form for each State vital statistics office.	132
Ynnnnnn.MATCH	MATCHING User Records (MATCH) – User records involved in matches. Same record format as submitted by the user.	100
Ynnnnnn.NOMATCH	NONMATCHING User Records (NOMATCH) – User records not involved in matches with any NDI records. Same record format as submitted by the user.	100
Ynnnnnn.REJECTS	REJECTED User Records (REJECTS) – User records rejected by the NDI edit program and not included in the NDI search. Same record format as submitted by the user.	100

* The *CAUSE* and *PRTCAUSE* files are only provided to NDI users who request the *NDI Plus* service to obtain cause of death codes.