## Linking Cancer Registry Data with the National Death Index (NDI): Information and Steps for Successful Linkage

NAACCR 2008-2009 Webinar Series

May 7, 2009

1

## Linking Cancer Registry Data with the National Death Index (NDI): Information and Steps for Successful Linkage

NAACCR 2008-2009 Webinar Series

May 7, 2009

Questions?

1

## Linking Cancer Registry Data with the National Death Index (NDI): Information and Steps for Successful Linkage

NAACCR 2008-2009 Webinar Series

May 7, 2009

Type questions here

1

## *Match a pet to a speaker and win a prize !!!*


4

## Webinar Objectives

Participants will:

- Become familiar with the process of conducting linkages between the NDI and cancer registry data
- Understand issues related to the use and re-release of the linked data

5

**NEED HANDS ON EXPERIENCE?**
**Attend the**
**1 day workshop, Friday, June 19, 2009**
**Following NAACCR Conference (San Diego)**

***

***Linking Cancer Registry Data with the National Death Index***

- If you plan to conduct NDI linkages, this workshop will give you the tools you need for a successful linkage
- Hands on experience with edits, files preparation, linkage results and clerical review using actual de-identified data
- You are encouraged (but not required) to bring a laptop.
- For more information on the workshop, please contact Hannah Weir by phone (770 488-30060) or email (hbw4@cdc.gov)

6

## Webinar Instructors

- Lyn Almon, Georgia Comprehensive Cancer Registry
- Robert Bilgrad, National Centers for Health Statistics, CDC
- Glenn Copeland, Michigan Cancer Surveillance Program
- Chris Johnson, Cancer Data Registry of Idaho
- David O'Brien, Alaska Cancer Registry
- Hannah K. Weir, DCPC, CDC
- Brad Wohler, Florida Cancer Data System
- Kevin Zhang, Macro International.

7

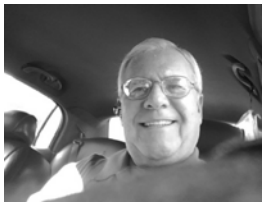
## Outline

| | | |
|---|---|---|
| 1. | Overview of the NDI | - Robert |
| 2. | NPCR-NDI application | - Hannah |
| 3. | The Case for NDI | - Chris |
| | Q & A | |
| 4. | CDC support for NDI Linkages | - Kevin |
| 5. | Preparing data for linkage | - David |
| 6. | Sending data to NDI | - Lyn / Chris |
| | Q & A | |
| 7. | Results from NDI | - Robert /Chris |
| 8. | Clerical review | - Glenn |
| 9. | Incorporation and Use of NDI data | - Brad |
| | Q & A | |

8

## Overview of the NDI (National Death Index)

Robert Bilgrad, MA, MPH
Special Assistant to the Director
Division of Vital Statistics
National Center for Health Statistics



9

## Why am I here?

- What is an NDI?
- Purpose
- Usage
- How it works
- Matching criteria
- NDI search results
- NDI *Plus*
- NDI charges

10

## Purpose of the NDI

- Identifies deceased study subjects
- Provides the following:
  - dates of death
  - states of death
  - death certificate numbers

13

## NDI PLUS

- Implemented in 1997
- Provides researchers with
  - Underlying cause of death codes
  - Multiple cause codes
  - ICD-9 and ICD-10 codes

14

## Coverage

- All 50 states, District of Columbia, NYC, Puerto Rico, & Virgin Islands
- 61 million NDI records
- All deaths from 1979-2006
- 2007 deaths May 2009??

15

## Studies using the NDI

- Clinical trials
- Post-marketing drug surveillance
- Medical/surgical treatment effectiveness
- Cancer, AIDS & other disease registries
- General population studies:
  - Census Bureau, NLMS (2.4 million records)
  - American Cancer Society (1 million records)
  - Dept. of Veterans Affairs (3 million records)

16

## NDI Website

www.cdc.gov/nchs/ndi.htm

17

## How to Use the NDI

- Submit NDI Application Form ***(done)***
- Allow 3 months for NDI adviser review **(done)**
- Submit records in NDI format
- Allow about 1 week for NDI results
- Assess NDI output files
- Identify true, questionable and false matches
- Purchase death certificates if necessary

18

## Data Items Used for NDI Searches

- First & last name, middle initial
- Father's surname (for females)
- Social Security Number
- Date of birth
- Sex

19

## Data Items for Assessing Matches

- State of birth -- (highly desirable)
- State of residence -- (last known)
- Race
- Marital status
- Age at death -- (for known decedents)

20

- Preparing user records
- Submitting records

21

## Seven Matching Criteria

1. Social Security Number

***Month and (+/-) 1 year of birth***

2. First name & last name
3. First initial, middle initial, & last name
4. First name & father's surname

***Birth Month and Day***

5. First name & last name
6. First initial, middle initial, & last name

22

## 7. For Women Only

Following items compared:
– Last name (on subject's record)
– Father's surname (on NDI record)

Following items must also agree:
– Month and year of birth
– First name

23

## NDI Retrieval Report

- Matching user record listed first
- Followed by one or more possible matches
  - State of death
  - Death certificate number
  - Date of death
- Shows which items agreed or disagreed

24

## Assessing the NDI Output (finding the right body)

- Many false NDI matches are generated
- Identify ***TRUE*** matches carefully
- Eliminate obvious ***FALSE*** matches
- Minimize remaining ***QUESTIONABLES***

25

## Submitting Test Records of KNOWN Decedents

- Expected TRUE matches: 92% - 98%
- Include dates of death
- TRUE matches: Assess items that agree
- FALSE matches: Assess items that disagree
- Apply assessment of known decedents to NDI match results for lost contact cases

26

## Factors Affecting NDI Matches

- Completeness and quality of information:
  - -- on study subjects
  - -- on state death records
- Effectiveness of NDI matching criteria
- Researcher's ability to assess NDI output

27

## NDI Output Files

- 11 files
- Text file formats
- CD-ROM
- Password protected using *PointSec*
- Express mailed within 1 week

28

## Confidentiality

NCHS assures users that their data will:

- Be held in strict confidence.
- Be used only to search the NDI file.
- Not be released to others without consent.
- Only be kept by NCHS for a short time.

29

## NDI User Fees

- NO CHARGE to cancer registries
- But still submit fee calculations
- Submit as many records as necessary
- All charges are covered by CDC and NCI

30

NDI Website

http://www.cdc.gov/nchs/ndi.htm

31


## NATIONAL DEATH INDEX

Robert Bilgrad, MA, MPH
Special Assistant to the Director
Division of Vital Statistics
National Center for Health Statistics
3311 Toledo Road, Room 7318
Hyattsville, MD 20782
Phone: 301-458-4101
E-mail: RBilgrad@cdc.gov
**********
NDI phone: 301-458-4444
NDI fax: 301-458-4034
NDI e-mail: ndi@cdc.gov

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

34
SAFER • HEALTHIER • PEOPLE™

## 2. NPCR-NDI Application

Hannah K. Weir, PhD
Division of Cancer Prevention and Control
Centers for Disease Control and Prevention

35

## NPCR-NDI Application

To obtain vital statistics information on patients who move out of state between the time of their diagnosis and death

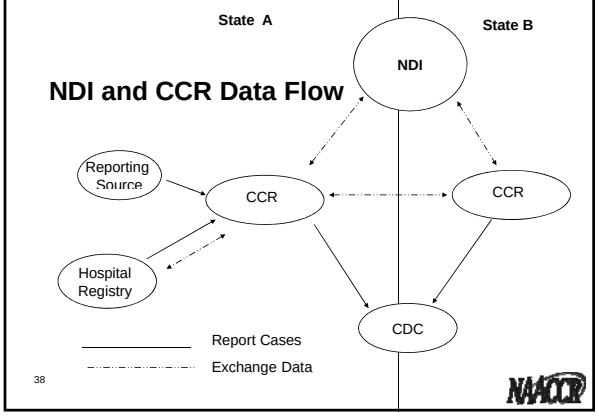
36

## NPCR-NDI Application

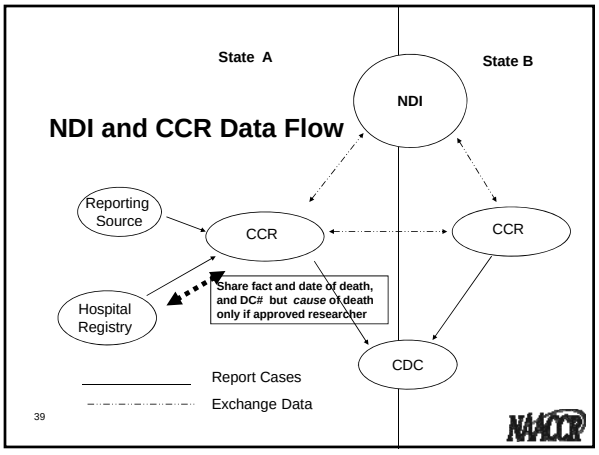
Spring 2006 –application approved for all cancer registries. The application includes the following:

- Release of NDI information (only fact and date of death but not causes) to hospitals for hospital follow up requirements
- Case sharing with other cancer registries to identify potential duplicate cases (fact, state and date of death, DC# but not cause)
- Release NDI information (fact, date and cause of death) to approved parties and to approved researchers
  - Registries must provide NDI with annual lists of researchers receiving identifiable death record information (i.e., researcher, organization, study title, date).

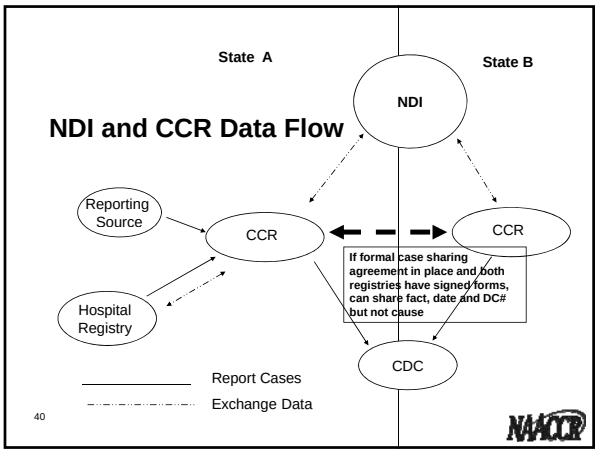
37

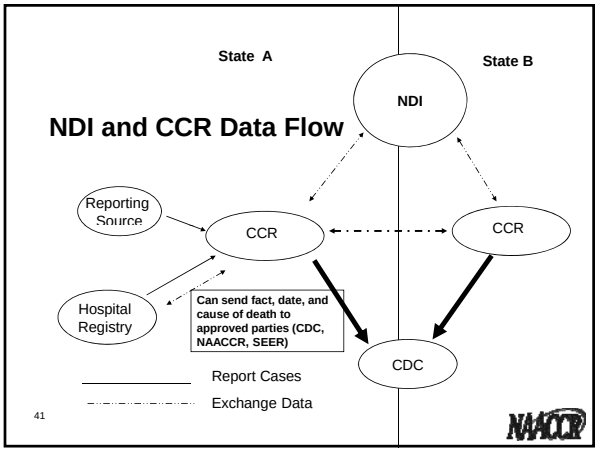


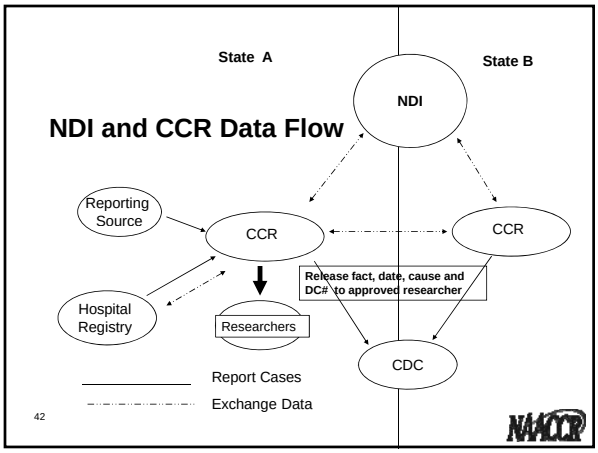

38




39


State A
State B
NDI
NDI and CCR Data Flow
Reporting Source
CCR
CCR
If formal case sharing agreement in place and both registries have signed forms, can share fact, date and DC# but not cause
Hospital Registry
CDC
Report Cases
Exchange Data
40
NAACCR


State A
State B
NDI
NDI and CCR Data Flow
Reporting Source
CCR
CCR
Hospital Registry
Can send fact, date, and cause of death to approved parties (CDC, NAACCR, SEER)
CDC
Report Cases
Exchange Data
41
NAACCR


State A
State B
NDI
NDI and CCR Data Flow
Reporting Source
CCR
CCR
Release fact, date, cause and DC# to approved researcher
Hospital Registry
Researchers
CDC
Report Cases
Exchange Data
42
NAACCR

**NAPHSIS Annual Meeting** **June 2006**
Met with NAPHSIS Executive Board to discus relationship between state cancer registries and vital statistics offices -- and arranged a reduced cost for cancer registries to use NDI

**Intra-Agency Agreement** **May 2007**
DCPC signed agreement with NCHS/NDI to cover the cost of NPCR-funded cancer registries to use NDI Plus services. A similar agreement is in place with SEER and their programs

## To Utilize NPCR-NDI Application

- Completed Supplemental NDI Confidentiality Agreement and Florida forms (do not complete NYC form)
- Sent forms to Hannah Weir => NCHS/NDI
- NDl issued (new) NDI number to each registry and send an approval package

44

## Suggested Data Requirement

- High quality incidence data (e.g., USCS publication criteria)
- Conduct linkage with state vital statistics records (all years, all causes of death)

45

## Contact Information

HBW4@CDC.GOV

770 488-3006

## 3. The Case for NDI

Chris Johnson
Epidemiologist
Cancer Data Registry
of Idaho

47

## The Case for NDI

- Promotes Collaboration
  - Promotes collaboration between states
    - States encouraged to share NDI results
  - Promotes collaboration between agencies
    - National level
    - State level
  - Reduce or eliminate "double counting"

48

## The Case for NDI

- Patient contact studies
  - Decedents will not be contacted
  - Spares families
  - Saves $$ for researchers

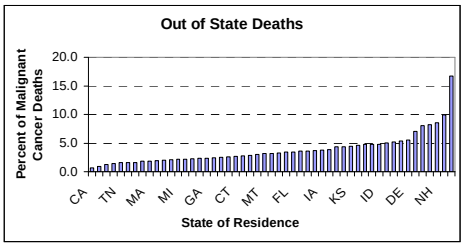
49

## Death Clearance Safety Net

- Late cases
  - Cancer registry
  - Vital Statistics
- Missed cases
  - Out of state

50

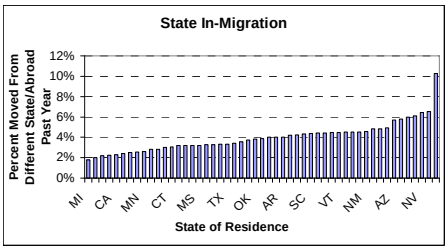
## Indirect Evidence




- 10 states had at least 5% of malignant cancer deaths occur out of state.

51

## Indirect Evidence




- From 2005-2007, 12 states had at least 5% of their population migrate in from another state or abroad.

52

## Indirect Evidence

- Americans very mobile
- Don't always die in their state of residence.
- Death missed during death ascertainment
  - ↓ event count for survival calculations (numerator)
  - ↑ increase in follow-up time (denominator).

53

## Direct Evidence

Benefits of Linkage to NDI in Longitudinal Study of Aging, James MK et al, J Aging Health. 1997 Aug; 9(3):298-315

- Examined agreement between re-interview and NDI
  - Substantially reduced bias due to lost of participants to follow up
  - NDI aided the study a great deal in capturing deaths among participants those living alone, below poverty index, without a proxy or did not supply phone number or did not own a home.

54

## Direct Evidence

A primer and comparative review of major US mortality databases, Cowper DC et al, Ann Epidemiology 2002 Oct;12(7) 462-8

- Out of all the national mortality ascertainment services, the NDI information demonstrated the highest sensitivity and it is currently the only source for cause of death information at the national level

55

## Florida Experience CONCORD Study

| Site | No. incident cases (1990-1994) | No. deaths identified by state files (i.e. DC present) | % deceased cases based on state data | No. deaths identified by NDI linkage (exclude sample of known decedents sent to NDI) | % deceased cases based on state (B) and national data (NDI Alone) |
|---|---|---|---|---|---|
| Prostate | **71,795** | **27,882** | 38.8% | **3,250** | 43.4% |
| Breast (female) | **61,936** | **19,534** | 31.5% | **2,087** | 34.9% |
| Colorectal | **58,885** | **33,417** | 56.7% | **3,013** | 61.9% |
| Male | **30,553** | **17,874** | 58.5% | **1,609** | 63.8% |
| Female | **28,332** | **15,543** | 54.9% | **1,404** | 59.8% |
| Total CONCORD | 192,616 | 80,833 | 42.0% | 8,350 | 46.3% |

56

## Florida Experience Preliminary Results

- Submitted 1+ million records
- 120,000+ records matched
- Death Safety Clearance Net
  - 9.6% matched that died in Florida
- 90.4% of deaths out of state
  - States with highest %
    - NY (9.7)
    - MI (7.4)
    - OH (6.6)
    - PE (5.4)
    - NJ (5.2)

57

***Match a pet to a speaker and win a prize !!!***


5

NAACCR

# Questions ???

NAACCR

## 4. CDC support for NDI Linkages

Kevin Zhang, PhD
Macro International

60

NAACCR

## Utilities Web Site for NDI

- Macro helps maintain a Utilities web site under the NPCR-CSS contract with CDC to facilitate the states NDI linkage efforts.
- NDI Utilities web site:
  - https://www.npcrcss.org/docserver
  - Select “utilities” and login
  - Click on “National Death Index (NDI)” box

61


62 Use your State user ID and password for access to the utilities web site and the document server for NPCR-CSS

## A Glance at the NDI Linkage Web Page


63 Use the NDI linkage utility programs on NPCR-CSS utilities web site to prepare data to send to NDI

## NDI Linkage Utility Programs

Three types of utility programs involving distinctive steps in the NDI linkage process are currently available for download from the NPCR-CS utilities web site:

- ***GenEdits plus for NDI:*** This edit program contains edit metafile specifically tailored for NDI linkage. Running this program on input data in NAACCR flat file format before extraction will help identify potential errors. Edits errors need to be resolved prior to data extraction.
- ***NDI extraction utility:*** A windows application that uses input file in NAACCR format and output a file in NDI-specified layout. This is a program for states who do not want to extract their data in SAS.
- ***Name frequency files:*** These are the files for post-linkage data processing. The zip file contains name frequency SAS files, V3 SAS algorithm and MS Access template to run the matching.

64

## Other Resources from Macro

- In addition to the utility programs shown above, Macro also helps CDC in distributing Social Security Death Index (SSDI) datasets to state central registries and other users.

-- We also developed a secure web-based data query system using the SSDI data.

-- States have the options of accessing the SSDI data via either the on-line query system or the batch mode by downloading the SSDI datasets from the secure document server.

65

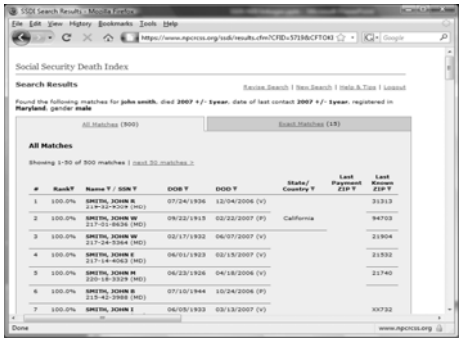
## SSDI Lookup: On-line Query System

- On-line query system can be used if you have only a small number of cases
- Only one name queried at a time
- NPCR secure web site: https://www.npcrcss.org/ssdi/login.cfm & needs user ID and password for access
- Contact NPCR-CSS Helpline at 301-572-0502 or email through: npcr-css@orcmacro.com.

66


CDC's SSDI on-line query system (secure site)

67



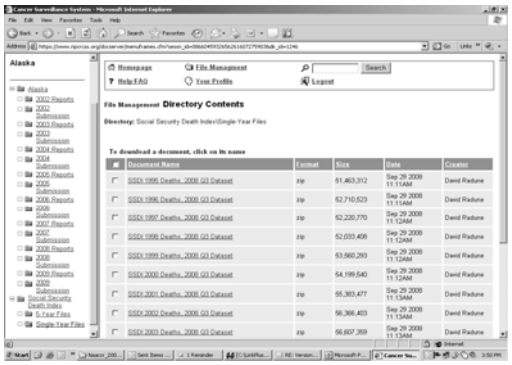

Results from on-line query system for **john smith**, died **2007 +/- 1year**, date of last contact **2007 +/- 1year**, registered in **Maryland**, gender **male**

68

## SSDI linkage: Batch Mode

- Batch mode linkage should be used for large number of registry cases
- SSDI data files can be downloaded from NPCR secure "docserver" web site: https://www.npcrcss.org/docserver/ (select "Upload Data" radio button at the login)
- Needs user ID and password for access (the same as for Call For Data)
- SSDI data files updated quarterly
- Use Link Plus or similar program for linkage

69




70 SSDI Single-Year Files on the NPCR-CSS docserver – download the SSDI file documentation FIRST (it is the last file on the list), it includes record layout

## 5. Preparing Data for NDI Linkage by First Doing SSDI Linkage

David O'Brien, PhD, GISP
Alaska Cancer Registry

71

## What is the SSDI?

- Social Security Death Index
- Database of all deceased Social Security Administration beneficiaries
- Data items: SSN, name, birth date, death date, state of residence, ZIP code last residence, ZIP code last SSA payment
- Does not contain cause of death
- Access by On-line Query System or Batch Mode

72

## Why Link with SSDI?

- Update registry case vital status more frequently than with NDI (esp. for survival analysis)
- More control over match determination than with NDI (can see details of matched pairs)
- Update registry case demographics (important for NDI)
- SSDI matches more likely to match NDI (to obtain Cause of Death)

73

## Preparing Access to SSDI in Batch Mode

- Install Link Plus
  http://www.cdc.gov/cancer/npcr/tools/registryplus/lp.htm
- Download all single-year SSDI files from NPCR "Doc Server"
  https://www.npcrcss.org/docserver/
- Export cases from registry database:
  - All live
  - Dead w/unk Cause of Death (7777 & 7797)
  - Dead w/unk SSN or DOB (incl. unk month or day)

74

## Run Edits on Registry Data

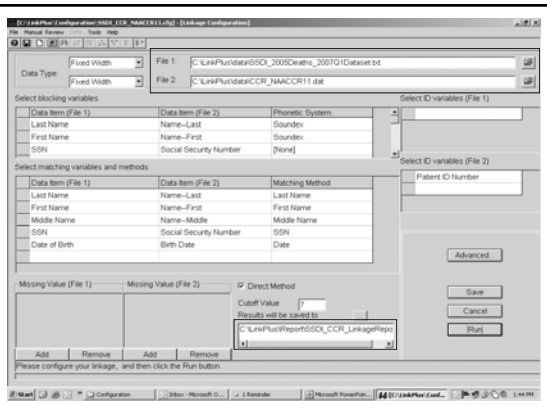
- Download GenEDITS Plus from NPCR Doc Server
- Metafile: NDI_v11_2.rmf
- Edit Set: NDI Edits
  - Includes many demographic edits (e.g., Name & SSN)
- Might be first time these edits ever run on registry data!
- Fix edits and re-export
- Run NPCR Inter Record edits

75

## Running Link Plus for SSDI Linkage

- Check for Link Plus files for SSDI linkage:
  - Configuration file: SSDI_CCR_NAACCR11.cfg
  - Record layout for SSDI: SSDI_Default.txt
  - Record layout for NAACCR v11: NAACCR11Default.txt
- Start Link Plus
- Open SSDI configuration file
- Re-establish all file names and paths
- Assignment of File 1 & 2 is important
  - File 1 = SSDI file (larger file)
  - File 2 = Registry file (smaller file)

76



**Re-establish file names and paths**

77

**Re-establish record layout file names and paths – click "View Data" to verify**

78

## Link Plus SSDI Config Settings

- Blocking variables:
  - Last Name (soundex)
  - First Name (soundex)
  - SSN
  - Birth Date
  - Zip code last residence (in SSDI file) / Addr Current--Postal Code (in Registry file)

79

## Link Plus SSDI Config Settings

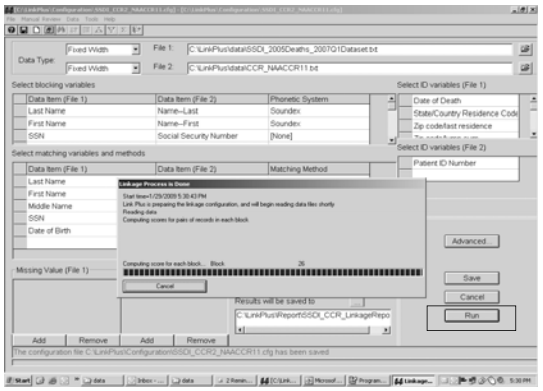
- Matching variables:
  - Last Name
  - First Name
  - Middle Name
  - SSN
  - Birth Date
- ID variables (for File 2 only):
  - Patient ID
- Use of ID variables affects program runtime

80

## Alaska-Specific Config Changes

- Added additional ID variables for File 1:
  - Date of Death
  - State/Country residence code
  - Zip code last residence
  - Zip code lump sum payment
- Changed cut-off from 7 to 10
  - For Alaska, most matches stopped around 15
  - For Alaska, 70% of matching report had scores between 7 and 10
- Might consider removing Zip Code and/or First Name as blocking variables to reduce program run-time

81

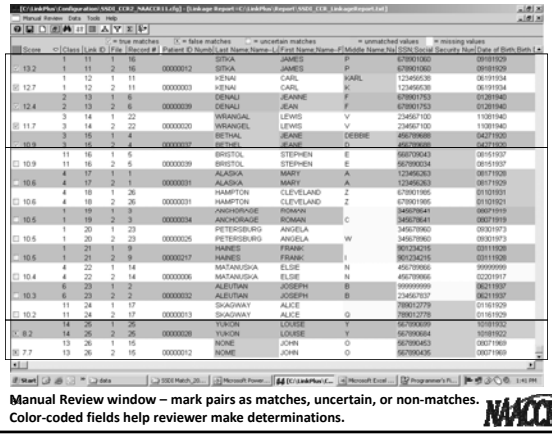


Click "Run" – Progress dialog box will appear

## Reviewing Match Results in Link Plus Manual Review Window

- Pairs are weighted by score
- Determine true matches, uncertain matches, and non-matches (automatically by score range, or manual selection)
- Fields are color-coded to show unmatched values and missing values
- Can hide ID fields because not in both files
- Can export separate files for true matches, uncertain matches, and non-matches

83



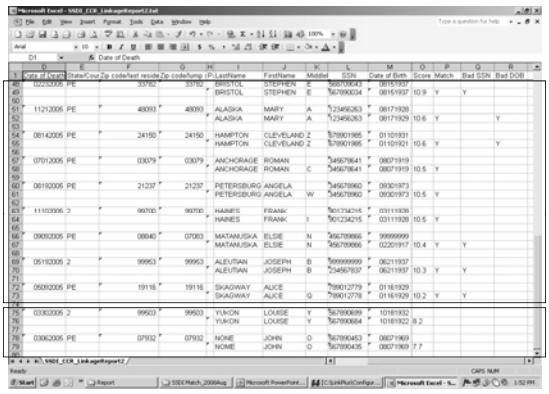
Yes Uncertain No

Manual Review window – mark pairs as matches, uncertain, or non-matches. Color-coded fields help reviewer make determinations.

## Match Results Review Process Used by Alaska (Overview)

- Import Link Plus linkage report into Excel (we don't use Manual Review window)
- Perform extensive research on uncertain matches to determine match status
- Correct registry DOB & SSN in Link Plus match report
- Link match report to registry data
  - Populate a "SSDI Link" non-NAACCR data item
  - Update corrected values of SSN and DOB
  - Update vital status-related data items

85



Uncertain

No

**Manual Review in Excel – mark matching pairs. Research unmatching DOB and SSN.**

## Match Results Review Process

- Very time consuming process for first-time match!
- Easier to do for future matches
- Alaska's first-time SSDI match stats (Aug 2008)
  - Approx 200 SSDI true matches per death year
  - 6.5% of all reportable cases matched to SSDI

87

## What If My Registry Can't Research Uncertain Matches?

- Can determine score range of just true matches
  - Update vital status in registry database
- Can create "alias records" for each uncertain match pair in which DOB, SSN, or Name differ

88

**Thanks very much!**

89

## 6. Sending Data to NDI

**Lyn Almon, MSPH, CTR**
**Georgia Center for Cancer Statistics**

90

## Coding Instructions for NDI Submission

- National Death Index User's Guide
  http://www.cdc.gov/nchs/r&d/ndi/ndi_usrsguide.htm
- Chapter II
  Preparing your records: Record layout and coding specifications

91

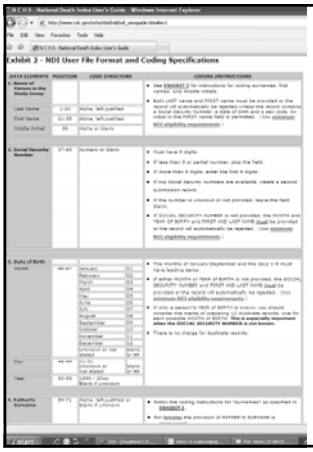



Exhibit 2- NDI User File Format

## NDI User Request Records

- Records for persons in cancer registry
  - Alive
  - Deceased with unknown cause of death
- Standard text file
- 100 characters per record

93

## Data Elements Needed for User File Creation

- Patient Id number
- NAACCR Record Version
- State of Residence
- Marital Status at DX
- Race 1
- Sex
- Birth Date
- Birthplace
- Dx Date
- Date of Last Contact
- Vital Status
- Cause of Death
- Last Name
- First Name
- Middle Name
- SSN
- Maiden Name

94

## Data Elements Contained in User Request Records

- Last Name
- First Name
- Middle Initial
- Social Security Number
- Birth Date
- Maiden Name
- Age at death
- Sex
- Race
- Marital Status
- State of residence
- Birthplace
- Patient ID Number (optional)
- Original or Alias indicator

95

## Each User Record Must Have:

- First and Last Name and SSN
  OR
- First and Last Name and Birth Month and Year
  OR
- SSN and DOB and Sex

96

## NAACCR Variables vs. NDI User Data File Variables

- Variables with different values
  - Race
  - Marital Status
  - Sex

97

## NPCR NDI Extraction Utility

- Input – NAACCR record format
- Output – NDI User Request record format
- Recodes necessary variables

98

## NPCR NDI Extraction Utility

- Server location: https://www.npcrcss.org
- Utilities page
- National Death Index
- Utilities – NDI Extraction Utility

99


## Extracted Data Sets

- Extracts at the patient level
- Multiple primaries- same patient id

  Original – Record with earliest dxdate
- Alias records

  Records with same patient ID, but difference in SSN, birthdate, or patient name

101

## Alias Records

- If unknown birth month – could input 12 records with each month of the year

  - Output 1 original record and 11 alias records
- If two records for same case with different names

  - Output 1 original record and 1 alias record

102

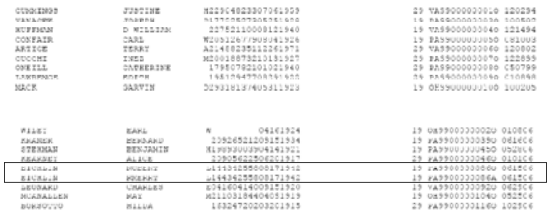
## Extraction Program Output

- Two datasets
  - Vital Status is alive (1) or missing
  - Vital Status is dead (0,4) and unknown cause of death
- Summary statistics

103

## Output Datasets

Two datasets: Alive or missing vital status and known dead/unknown cause of death



104

## Naming Conventions of Output

- Alive or missing – NDI Output File Name
  NDIUserRec.txt
- Dead – CauseOfDeathUnknown+NDI Output

  CauseOfDeathUnknownNDIUserRec.txt

105

## Summary Statistics

- Number of patient records with vital status live
- Number of duplicate records with vital status live
- Number of patient records with cause of death unknown
- Number of duplicate records with cause of death unknown

106

## Submission to NDI

- Encrypted/password protected output files from NPCR Extract Program
- CD-ROM
- Completed NDI Transmittal Form
- Completed Fee Worksheet
- Express mail CD with worksheet and transmittal Form
- Email password and mailing tracking number to NDI

107

## Fee Calculation

Chris Johnson

Cancer Data Registry of Idaho

108

## NDI User Fees

NDI

Effective October 1, 2004

**NATIONAL DEATH INDEX USER FEES**

The NDI is a self supporting service of the National Center for Health Statistics. NDI revenues are used primarily to cover the NDI's operating costs, especially the annual costs of purchasing files of death records from all of the state vital statistics offices.

| CHARGE PER STUDY SUBJECT* | |
|---|---|
| Vital status of each subject is UNKNOWN | $0.21 per subject -- per year of death searched |
| Subjects are all KNOWN to be deceased ** | $5.00 per decedent -- fixed fee |

The above charges are for NDI *Plus* services which also provide cause of death codes for the better matches. If your study only requires a *routine* NDI search (i.e., does not need cause of death codes), use $0.15 per subject (per year of death searched) for all of your subjects, including any subjects that are known to be deceased.

109

NAACCR

## NDI User Fees

* Charges are based on the number of *subjects*, not on the number of records submitted. Consequently, there is no charge for duplicate or alias records. To improve the matching effectiveness of your NDI search, you are encouraged to submit more than one record for those subjects having more than one first name, last name, father's surname, Social Security Number, or date of birth -- or for those subjects that appear to have nicknames.

** Whenever records of KNOWN decedents are submitted for a NDI *Plus* search, the deaths must have been identified *via sources other than the NDI* and must be submitted on a *separate file*. An exception to these NDI *Plus* charges for known decedents occurs whenever a NDI *Plus* user has already obtained copies of death certificates and simply wants to use NDI *Plus* to obtain the causes of death in coded form. The charges are only $2.50 whenever copies of certificates have already been obtained for each known decedent. (If certificates have only been obtained for some known decedents but not for others, the two groups of known decedents' records must be submitted on two *separate* files.)

110

NAACCR

## NDI User Fees

| SERVICE CHARGES *** | |
|---|---|
| Initial submission of user records . . . . . | $350.00 |
| Each subsequent submission . . . . . | $100.00 |

*** The service charge applies each time records are submitted. The $350 service charge is only for the *initial* submission of one or more *files* by a newly approved applicant. The $100 service charge is for each *subsequent* submission of one or more file.

***See Worksheet for Calculating NDI Charges.***

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention • National Center for Health Statistics
SAFER • HEALTHIER • PEOPLE™

CDC

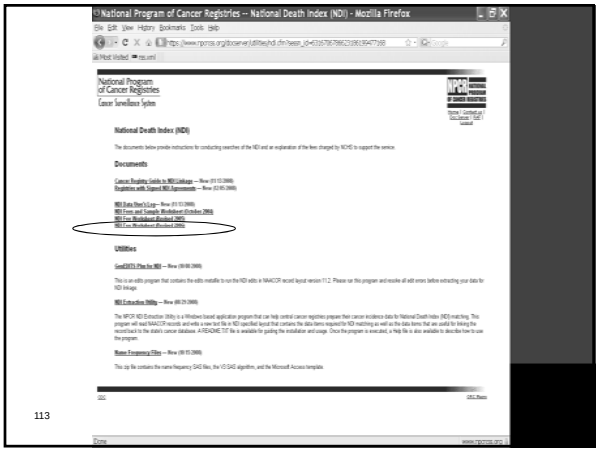
111

NAACCR

## NPCR/SEER NDI Fee Worksheet

- Where do you get it?
- How do you fill it out?
- How does this impact your processes?

112




113

## NPCR/SEER NDI Fee Worksheet

**NDI FEE WORKSHEET (for calculations through 2006)**

**NDI Plus Search: UNKNOWN Vital Status**

| First Year Searched | Last Year Searched | Number of Years Searched | | Number of Subjects | | NDI Fee* | | NDI Charge |
|---|---|---|---|---|---|---|---|---|
| 1995 | 2006 | 12 | X | 0 | X | $0.21 | = | $0.00 |
| 1996 | 2006 | 11 | X | 0 | X | $0.21 | = | $0.00 |
| 1997 | 2006 | 10 | X | 0 | X | $0.21 | = | $0.00 |
| 1998 | 2006 | 9 | X | 0 | X | $0.21 | = | $0.00 |
| 1999 | 2006 | 8 | X | 0 | X | $0.21 | = | $0.00 |
| 2000 | 2006 | 7 | X | 0 | X | $0.21 | = | $0.00 |
| 2001 | 2006 | 6 | X | 0 | X | $0.21 | = | $0.00 |
| 2002 | 2006 | 5 | X | 0 | X | $0.21 | = | $0.00 |
| 2003 | 2006 | 4 | X | 0 | X | $0.21 | = | $0.00 |
| 2004 | 2006 | 3 | X | 0 | X | $0.21 | = | $0.00 |
| 2005 | 2006 | 2 | X | 0 | X | $0.21 | = | $0.00 |
| 2006 | 2006 | 1 | X | 0 | X | $0.21 | = | $0.00 |
| | | Subtotal | | 0 | | | Subtotal | $0.00 |

* Use $0.15 per year if you are only doing a ROUTINE NDI search.

**NDI Plus Search: KNOWN Decedents**

| | Number of Subjects | | NDI Fee | | NDI Charge |
|---|---|---|---|---|---|
| Death certificates have NOT been obtained | 0 | X | $5.00 | = | $0.00 |
| Death certificates have been obtained | 0 | X | $2.50 | = | $0.00 |
| Subtotal | 0 | | | Subtotal | $0.00 |

Total Records: 0

**Service Charge:**

$350 for INITIAL search or $100 for REPEAT searches

Service Charge** $0.00

** This charge covers all files submitted at one time.

Total $0.00

114

## NPCR/SEER NDI Fee Worksheet

- **Preparing to send data to NDI:**
  - Complete the "NDI Fee Worksheet" to calculate the cost of the NDI linkages. First, calculate the number of eligible cases (i.e., vital status = "unknown") by diagnosis years. Enter the number of eligible cases by diagnosis year to calculate the number of required search years. On a separate line, enter the number of decedents for whom cause of death information is required.
  - Send an electronic copy of the worksheet to CDC or NCI, depending on which agency is covering the cost of your linkage.
    - Betsy Flagg eflagg@mail.nih.gov (NCI/SEER)
    - Hannah Weir (CDC/NPCR) (hbw4@cdc.gvo) (CDC/NPCR)

115

## NPCR/SEER NDI Fee Worksheet

- Submit the data files on a password protected CD(s) and send via express mail with the tracking number. ***Include a hard copy of the data transmittal form and the NDI Fee Worksheet***.
- When the package is mailed, alert NDI staff (NDI@cdc.gov) to expect a file and include the your unique registry NDI number, the transmittal form number (unique to each linkage) and the password for the CD. NDI tracks linkage request using the transmittal form number. At this time, NDI does not have a web server to receive data or return results.
- **REMEMBER: There is NO LIMIT to the number of records that can be sent to NDI.**

116

## NDI Search Years Available

- As of February 2009:
- 1979 – 2006 deaths

117

## NPCR/SEER NDI Fee Worksheet

- IF YOUR REGISTRY SUBMITTED A PATIENT TO NDI PREVIOUSLY, THAT MEANS YOU (CDC/NCI) DO NOT HAVE TO PAY FOR A SEARCH AGAINST THOSE YEARS AGAIN.
- FEES ARE CALCULATED FOR ONLY THE NEW YEARS OF DEATHS SEARCHED.
- THUS, YOU WILL NEED TO MATCH AGAINST ALL OF YOUR PREVIOUS LISTS OF CASES SENT TO NDI.

118

## Keeping Track of Previous NDI Submissions

- *** STEP 1 - READ IN PREVIOUS NDI SUBMISSION FILES AND CALCULATE PREVIOUS NDI YEARS SEARCHED ***;

| PreviousNDIFirstYear | PreviousNDILastYear | | | |
|---|---|---|---|---|
| | 1999 | 2003 | 2005 | Total |
| 1990 | 7,494 | 1,273 | 4,731 | 13,498 |
| 1996 | - | 175 | 449 | 624 |
| 1998 | - | - | 10,066 | 10,066 |
| Total | 7,494 | 1,448 | 15,246 | 24,188 |

119

## Central Cancer Registry Follow-Up

- *** STEP 2 - READ IN CURRENT CDRI DATABASE AND LOOK AT LAST FOLLOW-UP YEAR ***;

| LFUPYEAR | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 1970 | 396 | 0.3 | 396 | 0.3 |
| 1971 | 740 | 0.5 | 1,136 | 0.7 |
| 1972 | 923 | 0.6 | 2,059 | 1.3 |
| 1973 | 1,006 | 0.7 | 3,065 | 2.0 |
| 1974 | 1,204 | 0.8 | 4,269 | 2.8 |
| 1975 | 1,318 | 0.9 | 5,587 | 3.6 |
| 1976 | 1,483 | 1.0 | 7,070 | 4.6 |
| 1977 | 1,516 | 1.0 | 8,586 | 5.6 |
| 1978 | 1,565 | 1.0 | 10,151 | 6.6 |
| 1979 | 1,655 | 1.1 | 11,806 | 7.7 |
| 1980 | 1,895 | 1.2 | 13,701 | 8.9 |
| 1981 | 1,955 | 1.3 | 15,656 | 10.2 |
| 1982 | 2,128 | 1.4 | 17,784 | 11.5 |
| 1983 | 2,268 | 1.5 | 20,052 | 13.0 |
| 1984 | 2,336 | 1.5 | 22,388 | 14.5 |
| 1985 | 2,306 | 1.5 | 24,694 | 16.0 |
| 1986 | 2,517 | 1.6 | 27,211 | 17.7 |
| 1987 | 2,393 | 1.6 | 29,604 | 19.2 |
| 1988 | 2,542 | 1.7 | 32,146 | 20.9 |
| 1989 | 2,700 | 1.8 | 34,846 | 22.6 |

| LFUPYEAR | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 1990 | 2,854 | 1.9 | 37,700 | 24.5 |
| 1991 | 3,038 | 2.0 | 40,738 | 26.4 |
| 1992 | 3,149 | 2.0 | 43,887 | 28.5 |
| 1993 | 3,517 | 2.3 | 47,404 | 30.8 |
| 1994 | 3,699 | 2.4 | 51,103 | 33.2 |
| 1995 | 3,677 | 2.4 | 54,780 | 35.5 |
| 1996 | 3,893 | 2.5 | 58,673 | 38.1 |
| 1997 | 3,980 | 2.6 | 62,653 | 40.6 |
| 1998 | 4,357 | 2.8 | 67,010 | 43.5 |
| 1999 | 4,583 | 3.0 | 71,593 | 46.4 |
| 2000 | 4,844 | 3.1 | 76,437 | 49.6 |
| 2001 | 5,003 | 3.2 | 81,440 | 52.8 |
| 2002 | 5,309 | 3.4 | 86,749 | 56.3 |
| 2003 | 5,561 | 3.6 | 92,310 | 59.9 |
| 2004 | 6,103 | 4.0 | 98,413 | 63.8 |
| 2005 | 7,220 | 4.7 | 105,633 | 68.5 |
| 2006 | 7,990 | 5.2 | 113,623 | 73.7 |
| 2007 | 10,350 | 6.7 | 123,973 | 80.4 |
| 2008 | 29,504 | 19.1 | 153,477 | 99.6 |
| 2009 | 699 | 0.5 | 154,176 | 100.0 |

120

## Read Files from NPCR Extract

- By reading the files output from NPCR Extract and combining them with your historical NDI submission files, you can track each person and calculate fees accordingly.

121

## Cost Estimates - with Tracking

**NDI FEE WORKSHEET (for calculations through 2006)**

NDI Plus Search: UNKNOWN Vital Status

| First Year Searched | Last Year Searched | Number of Years Searched | | Number of Subjects | | NDI Fee* | | NDI Charge |
|---|---|---|---|---|---|---|---|---|
| 1979 | 2006 | 28 | X | 613 | X | $0.21 | = | $2,416.98 |
| 1980 | 2006 | 27 | X | 64 | X | $0.21 | = | $459.90 |
| 1981 | 2006 | 26 | X | 71 | X | $0.21 | = | $387.66 |
| 1982 | 2006 | 25 | X | 91 | X | $0.21 | = | $477.75 |
| 1983 | 2006 | 24 | X | 109 | X | $0.21 | = | $549.36 |
| 1984 | 2006 | 23 | X | 161 | X | $0.21 | = | $777.63 |
| 1985 | 2006 | 22 | X | 153 | X | $0.21 | = | $642.60 |
| 1986 | 2006 | 21 | X | 137 | X | $0.21 | = | $604.17 |
| 1987 | 2006 | 20 | X | 141 | X | $0.21 | = | $592.20 |
| 1988 | 2006 | 19 | X | 136 | X | $0.21 | = | $543.64 |
| 1989 | 2006 | 18 | X | 190 | X | $0.21 | = | $718.20 |
| 1990 | 2006 | 17 | X | 179 | X | $0.21 | = | $638.31 |
| 1991 | 2006 | 16 | X | 214 | X | $0.21 | = | $718.04 |
| 1992 | 2006 | 15 | X | 228 | X | $0.21 | = | $718.20 |
| 1993 | 2006 | 14 | X | 185 | X | $0.21 | = | $543.90 |
| 1994 | 2006 | 13 | X | 252 | X | $0.21 | = | $688.68 |
| 1995 | 2006 | 12 | X | 192 | X | $0.21 | = | $483.84 |
| 1996 | 2006 | 11 | X | 215 | X | $0.21 | = | $496.65 |
| 1997 | 2006 | 10 | X | 158 | X | $0.21 | = | $331.80 |
| 1998 | 2006 | 9 | X | 141 | X | $0.21 | = | $266.49 |
| 1999 | 2006 | 8 | X | 130 | X | $0.21 | = | $218.40 |
| 2000 | 2006 | 7 | X | 159 | X | $0.21 | = | $233.73 |
| 2001 | 2006 | 6 | X | 157 | X | $0.21 | = | $197.82 |
| 2002 | 2006 | 5 | X | 160 | X | $0.21 | = | $168.00 |
| 2003 | 2006 | 4 | X | 169 | X | $0.21 | = | $142.96 |
| 2004 | 2006 | 3 | X | 1198 | X | $0.21 | = | $754.74 |
| 2005 | 2006 | 2 | X | 505 | X | $0.21 | = | $212.10 |
| 2006 | 2006 | 1 | X | 1329 | X | $0.21 | = | $279.09 |
| | | | Subtotal | 17191 | | | Subtotal | $16,399.11 |

* Use $0.15 per year if you are only doing a ROUTINE NDI search.

NDI Plus Search: KNOWN Decedents

| | Number of Subjects | | NDI Fee | | NDI Charge |
|---|---|---|---|---|---|
| Death certificates have NOT been obtained | 4549 | X | $5.00 | = | $22,745.00 |
| Death certificates have been obtained | 0 | X | $5.00 | = | $0.00 |
| Subtotal | 4549 | | | Subtotal | $24,745.00 |

Total Records: 22140

Service Charge:

$350 for INITIAL search or $100 for REPEAT searches — Service Charge** $100.00

** This charge covers all files submitted at one time. — Total $50,243.11

122

## Cost Estimates – with no Tracking

**NDI FEE WORKSHEET (for calculations through 2006)**

NDI Plus Search: UNKNOWN Vital Status

| First Year Searched | Last Year Searched | Number of Years Searched | | Number of Subjects | | NDI Fee* | | NDI Charge |
|---|---|---|---|---|---|---|---|---|
| 1979 | 2006 | 28 | X | 613 | X | $0.21 | = | $2,416.98 |
| 1980 | 2006 | 27 | X | 64 | X | $0.21 | = | $459.90 |
| 1981 | 2006 | 26 | X | 71 | X | $0.21 | = | $387.66 |
| 1982 | 2006 | 25 | X | 91 | X | $0.21 | = | $477.75 |
| 1983 | 2006 | 24 | X | 109 | X | $0.21 | = | $549.36 |
| 1984 | 2006 | 23 | X | 161 | X | $0.21 | = | $777.63 |
| 1985 | 2006 | 22 | X | 153 | X | $0.21 | = | $642.60 |
| 1986 | 2006 | 21 | X | 137 | X | $0.21 | = | $604.17 |
| 1987 | 2006 | 20 | X | 141 | X | $0.21 | = | $592.20 |
| 1988 | 2006 | 19 | X | 136 | X | $0.21 | = | $543.64 |
| 1989 | 2006 | 18 | X | 190 | X | $0.21 | = | $718.20 |
| 1990 | 2006 | 17 | X | 179 | X | $0.21 | = | $638.31 |
| 1991 | 2006 | 16 | X | 214 | X | $0.21 | = | $718.04 |
| 1992 | 2006 | 15 | X | 228 | X | $0.21 | = | $718.20 |
| 1993 | 2006 | 14 | X | 185 | X | $0.21 | = | $543.90 |
| 1994 | 2006 | 13 | X | 252 | X | $0.21 | = | $688.68 |
| 1995 | 2006 | 12 | X | 192 | X | $0.21 | = | $483.84 |
| 1996 | 2006 | 11 | X | 215 | X | $0.21 | = | $496.65 |
| 1997 | 2006 | 10 | X | 158 | X | $0.21 | = | $331.80 |
| 1998 | 2006 | 9 | X | 141 | X | $0.21 | = | $266.49 |
| 1999 | 2006 | 8 | X | 130 | X | $0.21 | = | $218.40 |
| 2000 | 2006 | 7 | X | 159 | X | $0.21 | = | $233.73 |
| 2001 | 2006 | 6 | X | 157 | X | $0.21 | = | $197.82 |
| 2002 | 2006 | 5 | X | 160 | X | $0.21 | = | $168.00 |
| 2003 | 2006 | 4 | X | 169 | X | $0.21 | = | $142.96 |
| 2004 | 2006 | 3 | X | 1198 | X | $0.21 | = | $754.74 |
| 2005 | 2006 | 2 | X | 505 | X | $0.21 | = | $212.10 |
| 2006 | 2006 | 1 | X | 1329 | X | $0.21 | = | $279.09 |
| | | | Subtotal | 17191 | | | Subtotal | $25,198.11 |

* Use $0.15 per year if you are only doing a ROUTINE NDI search.

NDI Plus Search: KNOWN Decedents

| | Number of Subjects | | NDI Fee | | NDI Charge |
|---|---|---|---|---|---|
| Death certificates have NOT been obtained | 4549 | X | $5.00 | = | $24,745.00 |
| Death certificates have been obtained | 0 | X | $5.00 | = | $0.00 |
| Subtotal | 4549 | | | Subtotal | $24,745.00 |

Total Records: 22140

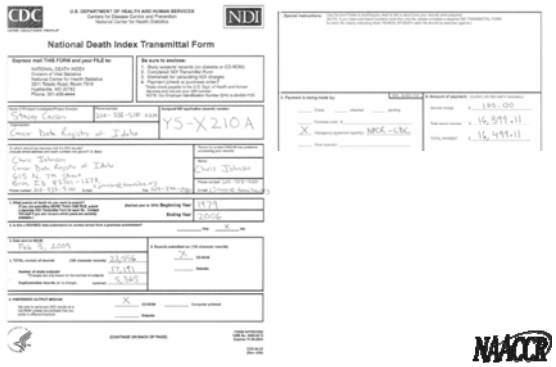
Service Charge:

$350 for INITIAL search or $100 for REPEAT searches — Service Charge** $100.00

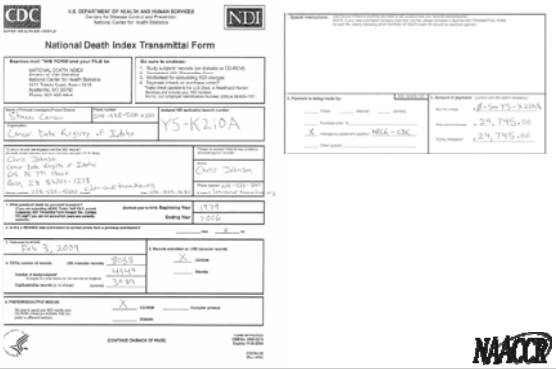
** This charge covers all files submitted at one time. — Total $50,043.11

123

## NDI Transmittal Forms




## NDI Transmittal Forms




## *Match a pet to a speaker and win a prize !!!*

*Questions ???*

## 7. Results from NDI

Robert Bilgrad
NCHS
Chris Johnson
Cancer Data Registry of Idaho

128

## NDI Output Files

- 11 files – NDI search results
- Text file formats
- ICD 9 and ICD 10 codes
- CD-ROM
- Expressed mailed within 1 week
- NDI User's Guide

129

## 11 NDI Output files

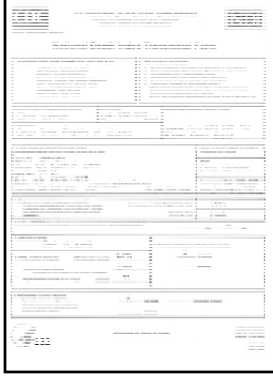
1. EDITS
2. SUMMARY
3. REPORT
4. COMPRESS
5. COMBINED
6. CAUSE
7. MATCH
8. PRTCAUSE
9. NOMATCH
10. REJECTS
11. REQFORMS

## NPCR NDI SAS Algorithm

- Example using training dataset.
- The data originated as real SSDI data and were sent for NDI processing. Noise was added to the data, which included deaths that occurred after 2006 (matched to NDI through 2006) so that there would be many non-matches. The death certificate numbers are fabricated. The causes of death are all 'ZZZZ'. The day component of the dates of death were changed.
- "This dataset is for demonstration of a hypothetical use of NDI. The data are from SSDI. NDI data have been obscured/made fictitious. Data are to be used within the organization only."

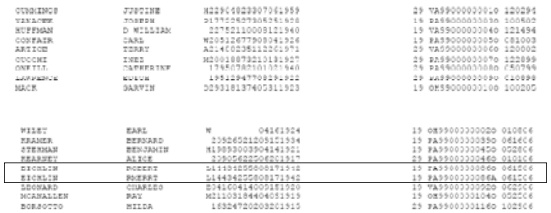
131

## Forms Sent to NDI



## Data Sent to NDI

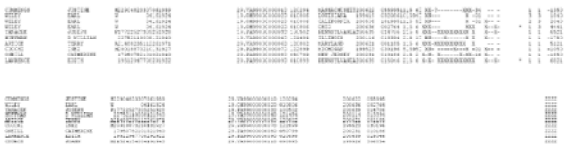
- Two datasets: unknown vital status and known dead.



133

## Data Returned from NDI

- Among the files:
- Y08NPCR.COMBINED
- Y08NPCR.CAUSE



134

## NDI Symbols

***Symbols used within the table:***

**X** = User's data item matched 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.

- (dash) = ***General usage*****:** Data item NOT provided on user record.

135

## NDI Symbols

***Exceptions for dashes*:**

**For SSN:** Specific digits of SSN did not agree.

**For SSN:** When SSN is not provided on user record, only one "-" is printed**.**

**For middle initial:** No middle initial was provided on user record; however, a middle initial was provided on NDI record.

**For "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.

***** = An asterisk in the left margin indicates that ALL data items provided on the user record matched exactly with the related items on the NDI record.

**N** = Names matched only on NYSIIS phonetic codes (New York State Identification and Intelligence System).

**I** = Only the first initials of the first names agreed on the user and NDI records.

**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 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 only existed on the NDI record, while a "?" signifies that the middle initial only existed on the user record.)

136

## NDI Symbols

**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 there is only one death certificate on file in that state.

**+01** = The birth year on the NDI record is one year more than birth year on the user record; for example, 1910 (NDI record) minus 1909 (user record) = +01.

**-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.

**±02..** = Difference between the years of birth on the NDI and user records. The four digit birth **thru** 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.

137

## NDI_SearchResults_CancerReg_Algorithm.sas

- This program is for use by NPCR and SEER registries that have conducted a National Death Index (NDI) file search, using the NDI Plus service to obtain cause of death information. The program processes the NDI '_____.cause' and NDI '_____.combined' file(s) and stratifies the possible matches into three groups:
- true matches,
- false matches,
- potential matches requiring manual review.

138

## NDI_SearchResults_CancerReg_Algorithm.sas

- Successful implementation of this algorithm assumes that each registry, when they submitted data to NDI, used the NPCR extraction utility which included the following NDI Optional User Data:
- ID (NAACCR item 20) in columns 82-89
- Original (O) or Alias (A) record flag in column 90
- Date of last contact or death (portion of NAACCR item 1750) in columns 92-97.

- Alternatively, registries could include this information without using the NPCR extraction utility or modify this program for user-specified Optional User Data.

139

## NDI_SearchResults_CancerReg_Algorithm.sas

- The date information is used to ensure that matches are not found when a person dies prior to the date of last known live follow-up. Should the date information not have been included in the data submitted to NDI, it will be necessary to merge your NDI results file(s) with your cancer registry database to obtain this information.

- The records returned from NDI in the "cause" files are a subset of the records included in the "combined" files. If you find a match with a record in the "combined" file that was not also included in the "cause" file, you will need to resubmit these to NDI to obtain cause of death information. Contact Robert Bilgrad at NDI if this occurs. The records that will need to be resubmitted will be output to the NDI_MATCHES_NOT_IN_CAUSE_FILE SAS dataset in the directory you specify in the SAS program.

140

## NDI_SearchResults_CancerReg_Algorithm.sas

- The SAS program requires the user to specify the name and location of two files and two libraries/folders/directories:

- *** THE PACKAGE OF RESULTS YOU RECEIVED FROM AN NDI FILE SEARCH INCLUDED ***;
- *** ONE OR MORE CDs OF THE SEARCH RESULTS. CHANGE THE FILENAME TO THE PATH ***;
- *** AND NAME OF YOUR NDI CAUSE RESULTS FILES. THESE WILL BE AMONG THE ***;
- *** FILES YOU UNZIPPED FROM THE NDI SEARCH RESULTS CDs. ***;
- filename cause 'v:\data\CANCER\DEATHS\NDImatch\NPCRNDI\NPCRNDINAACCRWebinarWorkshop\NDITrainingDataset\NDI_Results\Y08NPCR.CAUSE';

141

## NDI_SearchResults_CancerReg_Algorithm.sas

```
*** THE PACKAGE OF RESULTS YOU RECEIVED FROM AN NDI FILE SEARCH INCLUDED  ***;
*** ONE OR MORE CDs OF THE SEARCH RESULTS. CHANGE THE FILENAME TO THE PATH ***;
*** AND NAME OF YOUR NDI COMBINED RESULTS FILES.  THESE WILL BE AMONG THE  ***;
*** FILES YOU UNZIPPED FROM THE NDI SEARCH RESULTS CDs.                    ***;
filename combined
'V:\data\CANCER\DEATHS\NDImatch\NPCRNDI\NPCRNDINAACCRWebinarWorkshop\NDITrainingDataset\
NDI_Results\Y08NPCR.COMBINED';

*** CHANGE THE LIBNAME PATH TO THE DIRECTORY YOU WISH TO WRITE PERMANENT  ***;
*** SAS DATASETS TO.  TWO PERMANENT SAS DATASETS WILL BE WRITTEN, ONE FOR  ***;
*** RECORDS THAT THE ALGORITHM CODES AS BEING A TRUE MATCH, AND A SECOND   ***;
*** THAT INCLUDES ONLY THE RECORDS CODED AS POTENTIAL TRUE MATCH REQUIRING ***;
*** MANUAL REVIEW.  THESE FILES ARE NAMED:                                 ***;
***   NDI_TRUE_MATCH_RECORDS                                               ***;
***   NDI_MANUAL_REVIEW_RECORDS                                            ***;
*** AFTER THE MANUAL REVIEW, YOU CAN SET THE TWO DATASETS TOGETHER         ***;
*** AND UPDATE YOUR CANCER REGISTRY DATABASE WITH THE INFORMATION OBTAINED ***;
*** VIA THE NDI FILE SEARCH.                                               ***;
*** IF YOU HAVE ANY MATCHES FROM THE COMBINED FILE THAT ARE NOT ALSO IN    ***;
*** THE CAUSE FILE (FACT OF DEATH, BUT NO CAUSE OF DEATH INFORMATION),     ***;
*** THESE WILL BE PUT IN THE FILE NAMED:                                   ***;
***   NDI_MATCHES_NOT_IN_CAUSE_FILE                                        ***;
libname c
'V:\data\CANCER\DEATHS\NDImatch\NPCRNDI\NPCRNDINAACCRWebinarWorkshop\NDITrainingDataset\
SASAlgorithmEtc';
```

142

## NDI_SearchResults_CancerReg_Algorithm.sas

```
*** CHANGE THE LIBNAME PATH FOR THE DIRECTORY THAT CONTAINS THE SAS       ***;
*** NAME FREQUENCY DATASETS FROM NDI NAMES COUNT.
libname n 'V:\data\CANCER\DEATHS\NDImatch\NPCRNDI\NDINamesCount';
```

- The SAS program contains 8 steps:
- STEP 1 - READ IN CURRENT (1979-2005) NDI NAMES COUNT FILES
- STEP 2 - READ THE NDI COMBINED RESULTS FILE(S)
- STEP 3 - READ CAUSE OF DEATH INFORMATION FROM THE NDI CAUSE RESULTS FILE(S)

143

## NDI_SearchResults_CancerReg_Algorithm.sas

- STEP 4 - STATISTICS OF NDI MATCH FIELDS

***NDI SEARCH RESULTS***
***TABLES OF NDI MATCH FIELDS USED IN ALGORITHM***

*The FREQ Procedure*

| exact | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| * | 426 | 100.00 | 426 | 100.00 |

*Frequency Missing = 10123*

144

NDI_SearchResults_CancerReg_Algorithm.sas

| class | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| ssn7<seven items | 940 | 8.91 | 940 | 8.91 |
| missing ssn<eight items | 264 | 2.50 | 1204 | 11.41 |
| ssn doesnt match | 9345 | 88.59 | 10549 | 100.00 |

145

NDI_SearchResults_CancerReg_Algorithm.sas

| status | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| NDI Assume alive | 9634 | 91.33 | 9634 | 91.33 |
| NDI Assume dead | 915 | 8.67 | 10549 | 100.00 |

Frequency

Table of class by status

| class | status: NDI Assume alive | NDI Assume dead | Total |
|---|---|---|---|
| ssn7<seven items | 46 | 894 | 940 |
| missing ssn<eight items | 243 | 21 | 264 |
| ssn doesnt match | 9345 | 0 | 9345 |
| Total | 9634 | 915 | 10549 |

146

NDI_SearchResults_CancerReg_Algorithm.sas

***NDI SEARCH RESULTS***
***TABLES OF NDI MATCH FIELDS USED IN ALGORITHM***
***COUNT OF NUMBER OF NDI TRUE MATCHES PER CASE***

*The FREQ Procedure*

Frequency Count

| COUNT | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 1 | 839 | 95.78 | 839 | 95.78 |
| 2 | 36 | 4.11 | 875 | 99.89 |
| 4 | 1 | 0.11 | 876 | 100.00 |

147

NDI_SearchResults_CancerReg_Algorithm.sas

- STEP 5 - MERGE COMBINED RESULTS FILE WITH NAMES FILE IN ORDER TO OBTAIN PERCENT RANKS FOR EACH NAME, BEGIN THE NPCR NDI ALGORITHM

*NDI SEARCH RESULTS*
*FREQUENCY OF CALCULATED NAME FREQUENCY VARIABLE*

*The FREQ Procedure*

| rarename | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| 1 | 204 | 21.43 | 204 | 21.43 |
| 2 | 61 | 6.51 | 265 | 27.94 |
| 3 | 257 | 24.89 | 522 | 54.83 |
| 4 | 168 | 17.65 | 690 | 72.48 |
| 5 | 6 | 0.63 | 696 | 73.11 |
| 6 | 256 | 26.89 | 952 | 100.00 |

148

NDI_SearchResults_CancerReg_Algorithm.sas

- STEP 6 - USE NDI STATUS AND RESULTS OF NPCR NDI ALGORITHM TO DETERMINE IF RECORD IS A TRUE MATCH OR REQUIRES CLERICAL REVIEW
- STEP 7 - MERGE CAUSE OF DEATH INFORMATION FROM NDI "CAUSE" FILE
- STEP 8 - OUTPUT TRUE MATCH AND MANUAL REVIEW DATASETS AND PRODUCE RESULTS FOR NPCR ALGORITHM FOR INTERPRETING RESULTS OF NDI SEARCH

149

NDI_SearchResults_CancerReg_Algorithm.sas

*RESULTS OF NPCR ALGORITHM FOR INTERPRETING RESULTS OF NDI SEARCH*
*NOTE: RECORDS FOR MANUAL REVIEW MAY INCLUDE SEVERAL PER CANCER CASE*

*The FREQ Procedure*

| true | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
|---|---|---|---|---|
| NPCR True match | 849 | 93.71 | 849 | 93.71 |
| NPCR Manual review | 57 | 6.29 | 906 | 100.00 |

150

NDI_SearchResults_CancerReg_Algorithm.sas

*RESULTS OF NPCR ALGORITHM FOR INTERPRETING RESULTS OF NDI SEARCH*
*NOTE: RECORDS FOR MANUAL REVIEW MAY INCLUDE SEVERAL PER CANCER CASE*
*COUNT OF NUMBER OF MANUAL REVIEW MATCHES PER CASE*

*The FREQ Procedure*

| Frequency Count | | | | |
|---|---|---|---|---|
| COUNT | Frequency | Percent | Cumulative Frequency | Cumulative Percent |
| 1 | 55 | 98.21 | 55 | 98.21 |
| 2 | 1 | 1.79 | 56 | 100.00 |

151

NDI_SearchResults_CancerReg_Algorithm.sas

*RESULTS OF NPCR ALGORITHM FOR INTERPRETING RESULTS OF NDI SEARCH*
*TABLE OF STATUS (NDI) BY MATCH RESULTS (NPCR ALGORITHM)*

*The FREQ Procedure*

| Frequency | Table of status by true | | |
|---|---|---|---|
| | true | | |
| status | NPCR True match | NPCR Manual review | Total |
| NDI Assume alive | 2 | 28 | 30 |
| NDI Assume dead | 847 | 29 | 876 |
| Total | 849 | 57 | 906 |

152

NDI_SearchResults_CancerReg_Algorithm.sas

*RESULTS OF NPCR ALGORITHM FOR INTERPRETING RESULTS OF NDI SEARCH*
*TABLE OF PASS RESULTS (NPCR ALGORITHM)*

*The FREQ Procedure*

| Frequency | Table of pass by true | | |
|---|---|---|---|
| | true | | |
| pass | NPCR True match | NPCR Manual review | Total |
| 1 | 506 | 0 | 506 |
| 2 | 35 | 0 | 35 |
| 3.1 | 8 | 0 | 8 |
| 3.3 | 61 | 0 | 61 |
| 3.4 | 5 | 0 | 5 |
| 4.1 | 8 | 0 | 8 |
| 4.5 | 35 | 7 | 42 |
| 4.6 | 1 | 0 | 1 |
| 4.7 | 74 | 1 | 75 |
| 4.9 | 17 | 0 | 17 |
| 4.91 | 1 | 0 | 1 |
| 4.92 | 4 | 0 | 4 |
| 4.93 | 76 | 6 | 82 |
| 7 | 1 | 2 | 3 |
| 8 | 2 | 1 | 3 |
| 12 | 18 | 11 | 29 |
| 12.1 | 7 | 0 | 7 |
| 13.2 | 0 | 10 | 10 |
| 13.4 | 0 | 19 | 19 |
| Total | 849 | 57 | 906 |

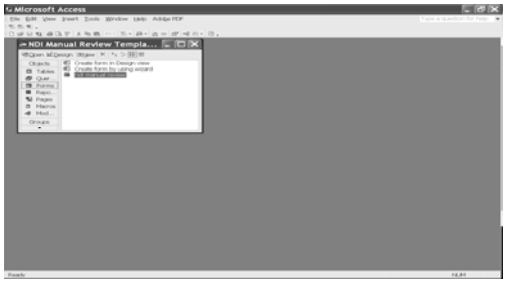
153

## NDI_SearchResults_CancerReg_Algorithm.sas

- AT THIS STAGE, YOU CAN USE SAS ACCESS OR DBMSCOPY OR ANOTHER DATA TRANSLATION TOOL TO CONVERT THE MANUAL REVIEW FILE TO MICROSOFT ACCESS FORMAT AND CONDUCT THE MANUAL REVIEW.
- THERE IS A BLANK ACCESS DATABASE WITH A FORM FOR USE IN THE MANUAL REVIEW PROCESS. THE BLANK DATABASE IS CALLED NDI Manual Review Template.mdb.
- YOU MUST SAVE YOUR DATA TO THE "NDI MANUAL REVIEW" TABLE FOR THE FORM TO WORK.

154

## NPCR Algorithm Manual Review

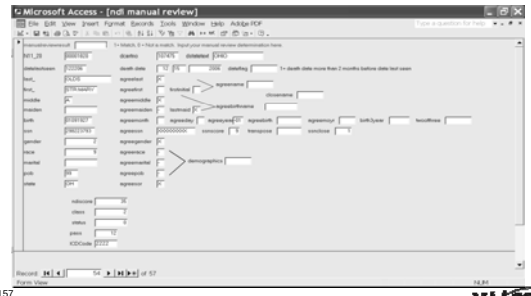



155

## NPCR Algorithm Manual Review


156

## NPCR Algorithm Manual Review




157

## 8. Clerical review

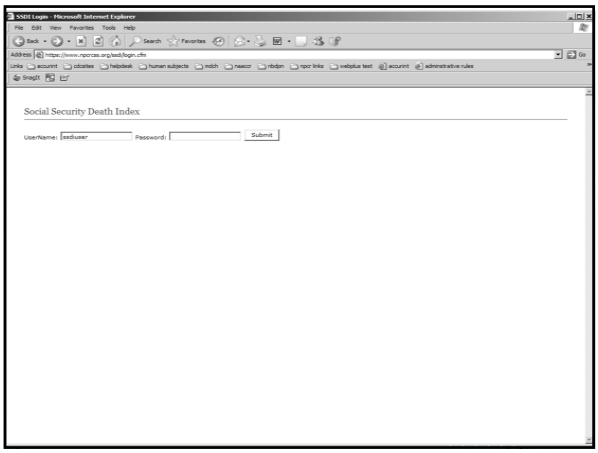
Glenn Copeland
Michigan Cancer
Surveillance Program

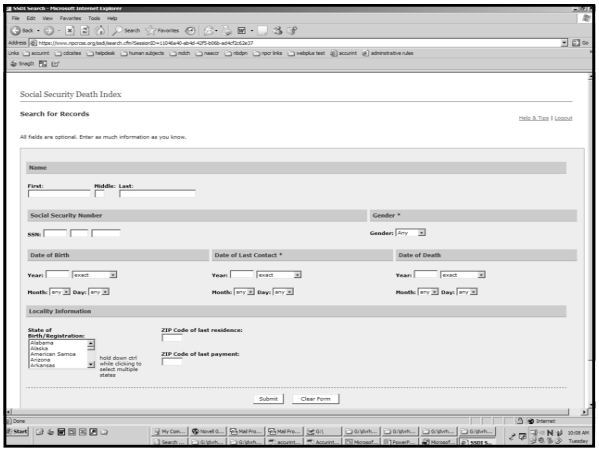
158

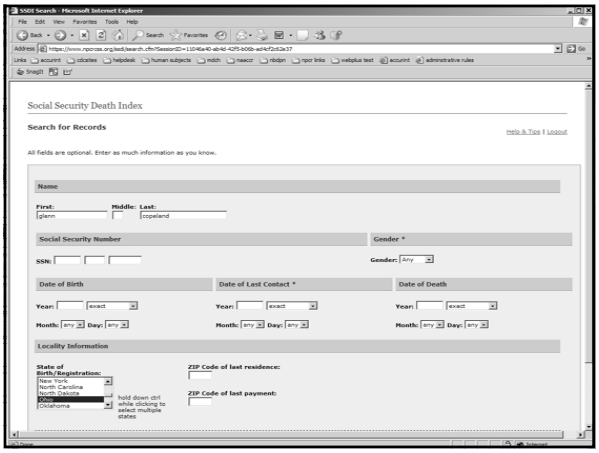
## Options for Verifying NDI Links

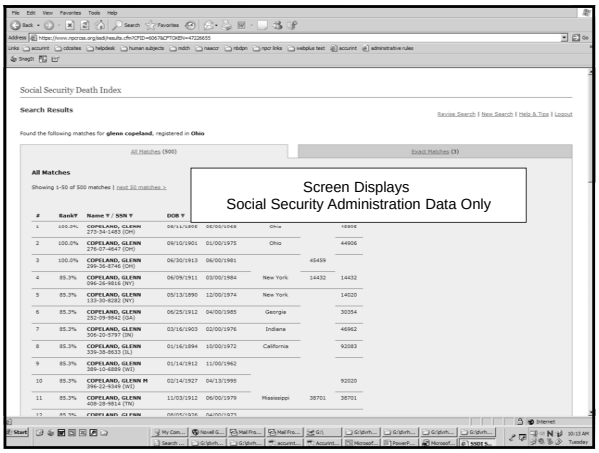
- Social Security Death Index
  - https://www.npcrcss.org/ssdi/
- Proprietary People Locator
  - www.accurint.com
- State Death Certificate
  - Contact individual states
  - NDI reference
    - Obtaining State Death Certificates

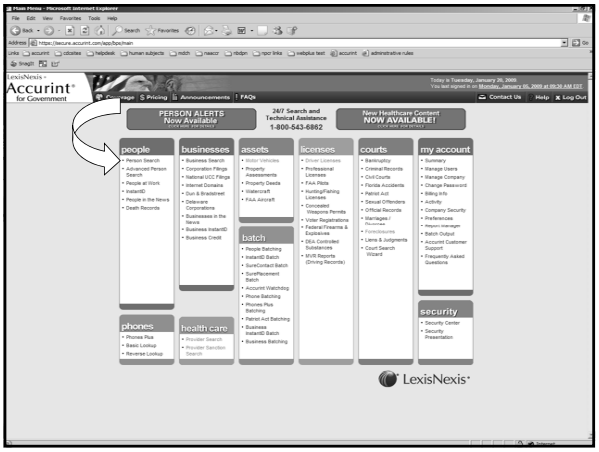
159

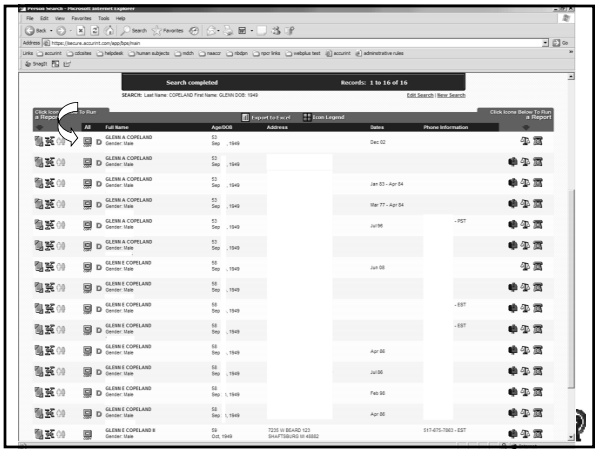

Social Security Death Index
UserName: ssdiuser
Password:
Submit

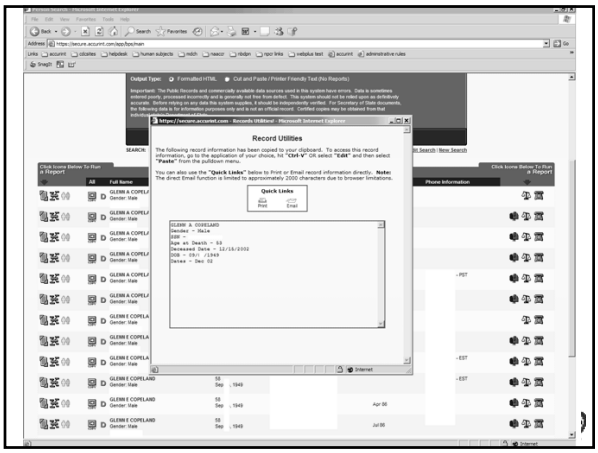

Social Security Death Index
Search for Records
Help & Tips | Logout
All fields are optional. Enter as much information as you know.
Name
First:
Middle:
Last:
Social Security Number
Gender *
SSN:
Gender:
Date of Birth
Date of Last Contact *
Date of Death
Year:
Month:
Day:
Locality Information
State of Birth/Registration:
ZIP Code of last residence:
ZIP Code of last payment:
hold down ctrl while clicking to select multiple states
Submit
Clear Form

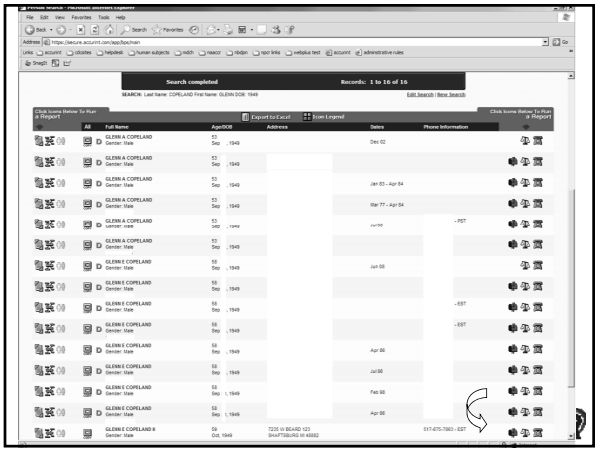

Social Security Death Index
Search for Records
Help & Tips | Logout
All fields are optional. Enter as much information as you know.
Name
First: glenn
Middle:
Last: copeland
Social Security Number
Gender *
SSN:
Gender:
Date of Birth
Date of Last Contact *
Date of Death
Year:
Month:
Day:
Locality Information
State of Birth/Registration:
New York
North Carolina
North Dakota
Ohio
Oklahoma
ZIP Code of last residence:
ZIP Code of last payment:
hold down ctrl while clicking to select multiple states


Social Security Death Index
Search Results
Found the following matches for glenn copeland, registered in Ohio
All Matches
Screen Displays
Social Security Administration Data Only


Accurint for Government
PERSON ALERTS
Now Available
24/7 Search and
Technical Assistance
1-800-543-6862
New Healthcare Content
NOW AVAILABLE!
people
businesses
assets
licenses
courts
my account
batch
phones
health care
security
LexisNexis


Search completed
Records: 1 to 16 of 16


Record Utilities
Quick Links


Search completed
Records: 1 to 16 of 16


Summary Report
Subject Information
AKAs
Indicators
Address Summary

## Accurint Fees

- Minimum $30.00 month
- $.25 per successful search
- $.50 address history
- $1 - $6 for detailed reports

169

## Obtaining Death Certificate Copies

- Obtaining State Death Certificates- 2008
  - Detailed information on state requirements
  - Provided with your NDI results

170

## 9. Incorporating and Using NDI data

Brad Wohler

Florida Cancer Data System

171

## Incorporating NDI Results

- Follow-up Source Central
- Vital Status
- Cause of Death
- Date of Last Contact
- Death Cert Number
- Place of Death (State)
- Date of last search

172

## Follow-Up Source Central

**FOLLOW-UP SOURCE CENTRAL**

| Alternate Name | Item # | Length | Source of Standard | Column # |
|---|---|---|---|---|
| | 1791 | 2 | NAACCR | 1397-1398 |

**Description**
This field is created by the central registry. It records the source from which the consolidated information was obtained on a patient's vital status and date of last contact. Follow-up Source Central would be updated when new or more reliable information becomes available. However, when the existing date of last contact/vital status is deemed to be more reliable than newly obtained information, then neither the date of last contact/vital status nor the follow-up source central would be changed.

**Rationale**
For central registries performing follow-up, this field could help evaluate the success rates of various methods of follow-up. When new follow-up information conflicts with the existing information, knowing the follow-up source can help resolve any discrepancies. Because follow-up information includes follow-up address and cancer status as well as date of last contact/vital status, and may come from different sources, it is important to note that Follow-up Source Central refers to the two fields, date of last contact and vital status.

| | |
|---|---|
| 04 | National Death Index (NDI) |

173

## Vital Status

**VITAL STATUS**

| Alternate Name | Item # | Length | Source of Standard | Column # |
|---|---|---|---|---|
| | 1760 | 1 | SEER/CoC | 1302-1302 |

**Description**
Vital status of the patient as of the date entered in Date of Last Contact [1750]. If the patient has multiple tumors, vital status should be the same for all tumors.

**Codes**

| | |
|---|---|
| 0 | Dead (COC) |
| 1 | Alive |
| 4 | Dead (SEER) |

174

## Cause of Death*

**CAUSE OF DEATH**

| Alternate Name | Item # | Length | Source of Standard | Column # |
|---|---|---|---|---|
| Underlying Cause of Death (SEER)<br>Underlying Cause of Death (ICD Code)<br>(pre-96 CoC) | 1910 | 4 | SEER | 1388-1391 |

**Description**
Official cause of death as coded from the death certificate in valid ICD-7, ICD-8, ICD-9, and ICD-10 codes.

**Rationale**
Cause of death is used for calculation of adjusted survival rates by the life table method. The adjustment corrects for deaths other than from the diagnosed cancer.

**Special codes in addition to ICD-7, ICD-8, ICD-9, and ICD-10 (refer to *SEER Program Code Manual* for additional instructions)**
0000 Patient alive at last contact
7777 State death certificate not available
7797 State death certificate available but underlying cause of death is not coded

*Note:* This data item is no longer supported by CoC (as of January 1, 2003).

**** Use care when sharing Cause of Death from NDI***

175

## Date of Last Contact

**DATE OF LAST CONTACT**

| Alternate Name | Item # | Length | Source of Standard | Column # |
|---|---|---|---|---|
| Date of Last Contact or Death (CoC)<br>Date of Last Follow-Up or of Death<br>(SEER) | 1750 | 8 | SEER/CoC | 1294-1301 |

**Description**
Date of last contact with the patient, or date of death. If the patient has multiple tumors, Date of Last Contact should be the same for all tumors. See page 95 for date format.

**Rationale**
Used for Date of Last Contact from active or passive follow-up. Used to record date of death.

176

## Death Certificate Number

**DC STATE FILE NUMBER**

| Alternate Name | Item # | Length | Source of Standard | Column # |
|---|---|---|---|---|
| | 2380 | 6 | State | 2278-2283 |

**Description**
Death certificate identification number as assigned by the vital statistics office in the place recorded in Place of Death [1940].

177

## Place of Death

**PLACE OF DEATH**

| Alternate Name | Item # | Length | Source of Standard | Column # |
|---|---|---|---|---|
| | 1940 | 3 | NPCR | 1394-1396 |

**Description**
State or country where the patient died and where certificate of death is filed.

**Rationale**
This field also helps carry out death clearance. When a hospital reports a place of death, the information can help in death certificate matching. It can also signal an out-of-state death for which the death certificate is to be requested.

**Codes in addition to geocodes**
997 Not applicable, patient alive
999 Place of death unknown

*Note:* See Appendix B for geocodes.

178

## Place of Death

- If death occurred outside of your state, the registry is encouraged to:
  - Share the dx record with the outside state provided
    - Current case sharing agreement in place
    - The state is a signatory to the NPCR NDI supplemental confidentiality agreement
- ***Do not share Cause of Death from NDI***

179

## Date of Last Search

- Date variable to store the date the record was last sent to NDI for match

180

## Secondary Data Release parties

- NPCR/NAACCR
- Signatory States
- Non-Signatory States
- Reporting Facilities
- Researchers

181

## NPCR/NAACCR

| | Can Be Shared |
|---|---|
| Fact of Death (Vital Status updated with NDI results) | Yes |
| Date of Death | Yes |
| Death Certificate Number | Yes |
| State of Death | Yes |
| Cause of Death | Yes |

182

## Other Signatory States*

| | Can Be Shared |
|---|---|
| Fact of Death (Vital Status updated with NDI results) | Yes |
| Date of Death | Yes |
| Death Certificate Number | Yes |
| State of Death | Yes |
| **Cause of Death** | **No** |

**Need to verify the Signatory States prior to sending out data , in order to be considered a signatory state, the state registry must have signed the NDI application.*
*Refer to the NPCR-CSS https://www.npcrcss.org/*
*Under the National Death Index link*

183 **Case sharing agreements in place with Signatory State*

## Other Signatory States

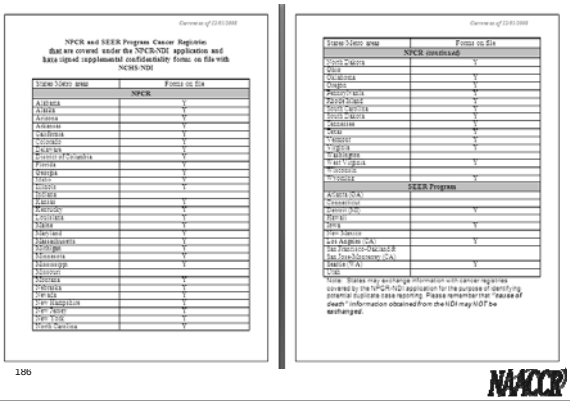
- Exchange of information (except cause) concerning cases associated with NDI is recommended and encouraged between signatory state registries

184

## Signatory States

– Patient dxed in state X and died in state Z
  - Residence at diagnosis State X
  - State X would receive NDI information and would also know that this person died in State Z
  - State X is encouraged to send the incident record to State Z
    – Help clear DCOs
    – Eliminates double counting

185



NPCR and SEER Program Cancer Registries that are covered under the NPCR-NDI application and have signed supplemental confidentiality forms on file with NCHS/NDI

| State/Metro area | Forms on file |
|---|---|
| **NPCR** | |
| Alabama | Y |
| Alaska | Y |
| Arizona | Y |
| Arkansas | Y |
| California | Y |
| Colorado | Y |
| Delaware | Y |
| District of Columbia | Y |
| Florida | Y |
| Georgia | Y |
| Hawaii | Y |
| Idaho | Y |
| Illinois | Y |
| Indiana | Y |
| Kentucky | Y |
| Louisiana | Y |
| Maine | Y |
| Maryland | Y |
| Massachusetts | Y |
| Michigan | Y |
| Minnesota | Y |
| Mississippi | Y |
| Missouri | Y |
| Montana | Y |
| Nebraska | Y |
| New Hampshire | Y |
| New Jersey | Y |
| New York | Y |
| North Carolina | Y |



| State/Metro area | Forms on file |
|---|---|
| **NPCR (continued)** | |
| North Dakota | Y |
| Ohio | Y |
| Oklahoma | Y |
| Oregon | Y |
| Pennsylvania | Y |
| Puerto Rico | Y |
| South Carolina | Y |
| South Dakota | Y |
| Tennessee | Y |
| Texas | Y |
| Utah | Y |
| Vermont | Y |
| Virginia | Y |
| Washington | Y |
| West Virginia | Y |
| Wisconsin | Y |
| Wyoming | Y |
| **SEER Program** | |
| Atlanta (GA) | |
| Connecticut | |
| Detroit (MI) | Y |
| Hawaii | Y |
| Iowa | Y |
| New Mexico | |
| Seattle (WA) | Y |
| Utah | |

Note: States may exchange information with cancer registries covered by the NPCR-NDI application for the purposes of identifying potential duplicate case reporting. Please remember that "cause of death" information obtained from the NDI may NOT be exchanged.

186

## Non-Signatory States**

| | Can Be Shared |
|---|---|
| Fact of Death (Vital Status updated with NDI results) | Yes |
| Date of Death | Yes |
| **Death Certificate Number** | **No** |
| **State of Death** | **No** |
| **Cause of Death** | **No** |

*** States that have not signed the NDI Application*

187

## Reporting Facilities

| | Can Be Shared |
|---|---|
| Fact of Death (Vital Status updated with NDI results) | Yes |
| Date of Death | Yes |
| **Death Certificate Number** | **No** |
| **State of Death** | **No** |
| **Cause of Death** | **No** |

188

## Researchers

| | Can Be Shared |
|---|---|
| Fact of Death (Vital Status updated with NDI results) | Yes |
| Date of Death | Yes |
| Death Certificate Number | Yes |
| State of Death | Yes |
| Cause of Death | Yes |

*Must maintain a log of all data releases containing NDI information*

189

## Data Release to Researchers

- NDI Advisory Group:
  - NDI data can be released to researchers as long as it is treated as confidential data and the researcher follows the state protocol for confidential data release
- IRB review recommended
- Cause of Death

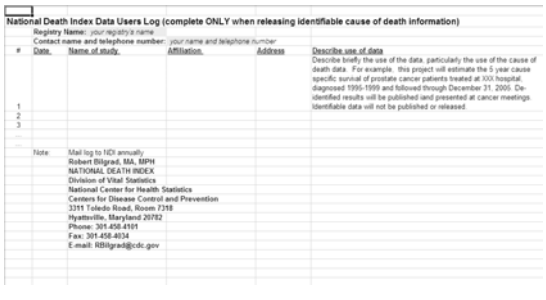
190

## Researcher Log

- At a minimum
  - State study ID number
  - Date data released
  - Name of study
  - Affiliation
  - Address
  - Describe use of data, particularly the use of the death data

191

## Researcher Log - Example



National Death Index Data Users Log (complete ONLY when releasing identifiable cause of death information)

Registry Name: *your registry's name*

Contact name and telephone number: *your name and telephone number*

| # | Date | Name of study | Affiliation | Address | Describe use of data |
|---|---|---|---|---|---|
| 1 | | | | | Describe briefly the use of the data, particularly the use of the cause of death data. For example, this project will estimate the 5 year cause specific survival of prostate cancer patients treated at XXX hospital, diagnosed 1995-1999 and followed through December 31, 2005. De-identified results will be published and presented at cancer meetings. Identifiable data will not be published or released. |
| 2 | | | | | |
| 3 | | | | | |

Note: Mail log to NDI annually:
Robert Bilgrad, MA, MPH
NATIONAL DEATH INDEX
Division of Vital Statistics
National Center for Health Statistics
Centers for Disease Control and Prevention
3311 Toledo Road, Room 7318
Hyattsville, Maryland 20782
Phone: 301.458.4101
Fax: 301.458.4034
E-mail: RBilgrad@cdc.gov

192

## Researcher Log

- The researcher log should be sent to Robert Bilgrad at NDI on an annual basis:
- **Robert Bilgrad, MA, MPH**
  **NATIONAL DEATH INDEX**
  **Division of Vital Statistics National Center for Health Statistics Centers for Disease Control and Prevention**
  **3311 Toledo Road, Room 7318**
  **Hyattsville, Maryland 20782**

  **Phone: 301-458-4101**
  **Fax: 301-458-4034**
  **E-mail: RBilgrad@cdc.gov**

193

## Where to get help

**ATTEND THE NDI WORKSHOP June 19th**
**NAACCR Meeting, San Diego**

NPCR-CSS Website (NDI page)

NAACCR Mentorship Program

Mentor registries???

194

## Questions ???

## *Match a pet to a speaker and win a prize !!!*


## Thank you for participating in today's webinar!

- The next webinar is scheduled for ***6/11/2009***.
  - The topic is '***Collecting Cancer Data: Prostate***'.
- Forward questions from today's webinar to Shannon or Jim.

**Shannon Vann**
svann@naaccr.org or 217-698-0800 X9
**Jim Hofferkamp**
jhofferkamp@naaccr.org or 217-698-0800 X5