

VIReC Technical Report 2:

VA-NDI Mortality Data Merge Project

Noreen Arnold, MS, Min-Woong Sohn, PhD,
Charles Maynard, PhD, and Denise M. Hynes, PhD, RN

April 9, 2006



VA Information Resource Center
Edward Hines, Jr. VA Hospital (151V)
P.O. Box 5000 • Hines, IL 60141
Ph: 708-202-2413 • Fax: 708-202-2415
virec@va.gov • www.virec.research.va.gov

Contents

| | |
|---|----|
| I. Executive Summary | 4 |
| Objectives | 4 |
| Background | 4 |
| Methods | 5 |
| Results | 5 |
| Recommendations | 6 |
| II. Development of the Study Cohort..... | 6 |
| Table 1: Number of Veterans Included in the Study..... | 7 |
| III. Identification of Dates of Death from Sources Available Within the VA | 7 |
| IV. Selection of Samples for Routine NDI Searches and SSA Epidemiological Vital Status Searches..... | 9 |
| Table 2: Random Samples for NDI and SSA Searches | 10 |
| V. Analysis of the NDI and SSA Search Results | 11 |
| Multiple Sources with the Same Date of Death (Table 2, Row 1)..... | 11 |
| Multiple Sources with Differing Dates of Death (Table 2, Row 2)..... | 12 |
| Differing month or day of death (Table 2, Row 2a.)..... | 12 |
| Differing year of death (Table, 2 Row 2b.) | 13 |
| Only One Source with a Date of Death (Table 2, Row 3)..... | 14 |
| Presumed Living (Table 2, Row 4)..... | 15 |
| Unknown Status (Table 2, Row 5)..... | 15 |
| Activity After Death (Table 2, Row 6)..... | 16 |
| Overall Sample (Table 2, Row 7) | 17 |
| Summary of Our Analyses | 18 |
| Limitations..... | 18 |
| Table 3: NDI Search Results from Cases with Dates of Death in Multiple Sources That Match..... | 20 |
| Table 4: NDI Search Results from Cases with Dates of Death in Multiple Sources That Differ Across the Sources..... | 21 |
| Table 5: Agreement Rates with NDI Date of Death for Cases with Dates of Death in Multiple Sources That Differ in Month or Day across the Sources..... | 22 |
| Table 6: Agreement Rates with NDI Date of Death for Cases with Dates of Death in Multiple Sources That Differ in Year across the Sources | 23 |
| Table 7: NDI Search Results from Cases with a Date of Death Only in One Source..... | 24 |
| Table 8: NDI Search Results for Veterans with an Unknown Status..... | 25 |
| Table 9: SSA Search Results for Veterans with an Unknown Status | 26 |
| Table 10: NDI Results for Veterans with Activity after Death..... | 27 |
| Table 11: Agreement Rates with NDI Date of Death for Cases with Activity after Death...28 | 28 |
| Table 12: Overall Sample Categories and NDI Deaths Identified..... | 29 |
| Table 13: Comparison of Deaths in the Four Sources (BIRLS-DF, SSA, Medicare, MSID) Available Within the VA and the NDI..... | 29 |
| Table 14: NDI Search Results for the Overall Sample..... | 30 |
| Appendix A: Files Used to Create the Study Cohort..... | 31 |
| Appendix B: Files Used to Obtain Demographics for an SSN..... | 32 |

Acknowledgements

This work was supported by the Health Services Research and Development Service funding for the VA-NDI Mortality Data Merge Project (SDR 03-157, Min-Woong Sohn, PI and Denise M. Hynes, Co-PI) and funding for the VA Information Resource Center (SDR 98-004, Denise M. Hynes, PI). The opinions expressed in this document are those of the authors and do not necessarily represent the opinions of the Department of Veterans Affairs.

Patricia A. Murphy, M.S., Kristin Koelling, M.P.H., and Cynthia Padera edited this report.

Suggested Citation

Arnold N, Sohn M, Maynard C, Hynes DM. VIREC Technical Report 2: VA-NDI Mortality Data Merge Project. Edward Hines, Jr. VA Hospital, Hines, IL: VA Information Resource Center, April 9, 2006.

VA-NDI Mortality Data Merge Project Technical Report

I. Executive Summary

Objectives

The Veterans Benefits Administration's (VBA) Beneficiary Identification Records Locator Subsystem (BIRLS) Death File is the main internal source utilized by Department of Veterans Affairs (VA) researchers to ascertain deaths in their study cohorts.¹⁻⁴ Anecdotal reporting suggests that VA researchers use other sources to augment BIRLS death data due to their concerns regarding its completeness and accuracy. Based on interest within the VA research community to acquire death dates from alternative sources including the National Center for Health Statistics' National Death Index (NDI), previous research examining the sensitivity of BIRLS death dates⁵, and our recent comparison of death dates in the BIRLS Death File (BIRLS-DF) to death dates provided by Medicare, this feasibility study was funded and undertaken to more accurately determine the costs and benefits associated with use of alternative sources for death ascertainment and to develop a strategy for the possible development of a central database or death registry that could be used by VA researchers for death ascertainment. The specific aims of this study were:

1. Analyze and compare the four sources for death data currently available within the VA: Medicare Vital Status Files, Social Security Administration Death Master File⁵⁻⁷, BIRLS Death File and VHA Medical SAS^{®*} Inpatient Datasets^{8,9}.
2. Determine the benefit of using the NDI to identify additional deaths and validate deaths in other sources.
3. Develop a methodology to use these sources most cost effectively for death ascertainment in research.
4. Propose a strategy to maintain a database of death date information and provide cost estimates for annual maintenance and updates.

This report details the methods used to accomplish these aims, the results of our analyses, and our recommendations regarding development of a death registry.

Background

The National Death Index is considered the gold standard for death ascertainment and was developed specifically for use by researchers.¹⁰ The NDI contains information from death certificates collected since 1979 by all fifty states, the District of Columbia, Puerto Rico and the Virgin Islands. The National Center for Health Statistics (NCHS) maintains the NDI. The current fees for using the NDI are \$0.15 per year searched per record, plus a flat \$350.00 charge per request submitted. The VA has served more than 10 million veterans since 1997. To annually request an NDI search for all of these veterans to ascertain deaths would cost over \$1.5 million. But if veterans for whom the VA has lost contact could be identified, the cost of using

*SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

the NDI could be reduced drastically by only requesting NDI searches for those veterans. Such judicious use of the NDI could result in cost effective death ascertainment.

Within the VA there are three sources for death ascertainment data that can be used to augment BIRLS death data: the Medicare Vital Status File, the Social Security Administration's Death Master File (SSA DMF), and the VHA Medical SAS Inpatient Datasets (MSID). These sources are described in detail in Section III. We combined the death data from the BIRLS Death File, the SSA DMF, the MSID and the Medicare Vital Status File and assessed the costs and benefits of using an NDI search to validate this pooled death data and identify additional deaths.

Methods

The first step was the development of a file of veterans "known to the VA" at any point in fiscal years 1997-2002, i.e., veterans who had enrolled in the Veterans Health Administration (VHA), received compensation or pension benefits, or utilized the VA Healthcare System. We identified a cohort of 8.6 million veterans fitting these parameters. Death data for this cohort were obtained from the four sources available within the VA. (See Section II) Veterans were then grouped into the following categories: deceased, presumed living and unknown status. (See Section IV) Stratified random samples were selected from the three categories for an NDI Search and for some samples an SSA Epidemiological Vital Status Search to identify additional deaths in the presumed living and unknown status categories and to validate the dates of death in the pooled death data. (See Section IV) The SSA Epidemiological search provides a presumption of living status that was used to validate our criteria for categorizing a veteran as presumed living. The search results were analyzed to determine the benefits of using the NDI. (See Section V)

Results

The following are the key results and conclusions from this feasibility study:

- The combined use of the BIRLS Death File, the SSA Death Master File, the Medicare Vital Status File and the MSID compares very favorably with the results of the NDI search. Thus, the cost of using the NDI under the current fee structure is not justified.
- The methodology developed to categorize veterans as presumed living or unknown status compares very favorably to the SSA Epidemiological Vital Status Search. Thus, adding a "presumed living" status to a Veteran's Death Registry based on information available within the VA (i.e., VHA usage dates, Medicare usage dates and enrollment, and dates of receipt of compensation and pension benefits) would provide valuable information to researchers.
- Use of external sources (SSA DMF and Medicare Vital Status File) increased the number of deaths identified by 22.4% over the use of only internal sources (BIRLS Death File and the MSID) for the years 1999-2002. This percentage increased from 20.5% in 1999 to 25.6% in 2002.
- The most significant data quality issue regarding death ascertainment using the four sources available within the VA is inaccurate recording of social security numbers or demographics. We estimate that about 2% of veteran records may be affected by these issues. The social security number was used to link veteran records across the four

sources, and date of birth and sex were used to determine if the linked records belonged to the same person.

- Death ascertainment in two small subgroups benefited significantly from an NDI search: veterans 65 and older classified with an unknown status and veterans with a date of death in only the BIRLS Death File or the MSID with activity more than a year after this date of death. Because these subgroups are such a small percentage of the overall population of veterans known to the VA, building the infrastructure to perform ongoing NDI searches for these veterans would not be cost effective. If researchers have significant numbers of veterans fall within these subgroups they might consider using the NDI.

Recommendations

Based on our results, we propose that a Veteran's Death Registry combining the death date information from the BIRLS Death File, the MSID, the Medicare Vital Status File, and the SSA Death Master File be developed. We also propose that the Registry include information about activity within the VA and in Medicare that could be used to establish a presumed living status.

The following sections describe the methods and results of our analyses in more detail.

II. Development of the Study Cohort

We identified veterans by applying the same methodology used annually to create a "finder" file of veterans for the VA Medicare Data Merge Initiative (SDR 02-237). This methodology identifies veterans who were VHA users, VHA enrollees, or recipients of compensation or pension benefits. Appendix A contains a list of the files used to compile this cohort of veterans. There were 9.4 million social security numbers (SSN) identified in these sources that were sent to the Center for Medicare and Medicaid Services (CMS) to obtain Medicare data.

Demographic data, including gender, date of birth and veteran status, were obtained from files listed in Appendix B. We found that the demographics for some SSNs were not consistent across all sources. This inconsistency may be due to the use of a veteran's SSN by more than one person or the inaccurate recording of the veteran's SSN or demographics. Thus, there may be several individuals in the system using the same SSN or individuals with inaccurately recorded SSNs or demographics. We used the most commonly occurring date of birth and gender. If the SSN was identified in any file as belonging to a veteran we considered the SSN to be that of a veteran.

Because the SSN and demographics were used to link veteran records across the source files, SSNs were removed from the study cohort if they were obviously invalid (i.e., the first three digits of the SSN were all zeroes or were greater than 772 which is outside the issued range¹¹) or if no date of birth was found in any source in Appendix B. SSNs were also removed if the SSN was not identified as belonging to a veteran in any source or if the SSN did not have an updated Medicare Vital Status record. The Medicare Vital Status File is received annually by the VA from CMS, and the annual file received in February 2004 was missing records for a small number of SSNs. The problem was reported to CMS. We eliminated these SSNs to reduce

bias that might be introduced due to non- receipt of an updated record. Table 1 provides statistics on SSNs removed from the study cohort:

Table 1: Number of Veterans Included in the Study

| Description | Number of SSNs (Veterans) | Percent of Original Cohort |
|---|----------------------------------|-----------------------------------|
| Original Study Cohort | 9,422,337 | 100.0% |
| Less invalid SSNs | 10,940 | 0.1% |
| Less SSNs with no date of birth | 27,314 | 0.3% |
| Less SSNs not identified as veterans | 725,557 | 7.7% |
| Less SSNs without an updated Medicare Vital Status record | 53,943 | 0.6% |
| Total SSNs remaining in the cohort | 8,604,583 | 91.3% |

There was a significantly larger percentage of females among those with SSNs that could not be identified as belonging to a veteran than in the remaining cohort (71.5% of those not identified as veterans were women; 4.7% of veterans were women). We suspect these individuals may be non-veteran spouses. There was also a significantly larger percentage of females among those without an updated Medicare Vital Status record than in the remaining cohort (19.4% of those without an updated record were women; 4.7% of those with an updated record were women). Because CMS has not responded to our inquiries about these records, there is no explanation for this finding.

III. Identification of Dates of Death from Sources Available Within the VA

Death dates were obtained from the following four sources:

1. the Beneficiary Identification and Records Locator Subsystem Death File (BIRLS-DF),
2. the Medical SAS Inpatient Datasets (MSID),
3. the Medicare Vital Status File, and
4. the Social Security Administration Death Master File (SSA DMF).

The **BIRLS Death File** is an extract from BIRLS and is updated monthly. BIRLS is updated with a date of death when the VBA receives notification that a veteran has died. Notification may come from family members applying for death benefits, VHA hospitals, or the VA National Cemetery Administration. The VBA also has a monthly, automated process to identify deaths by matching the Compensation and Pension Master File with the SSA Death File. Regional Offices receive these notifications and update BIRLS with the date of death manually. Only if the veteran is receiving benefits will the notice of death be validated to prevent termination of benefits if the notice is incorrect.* In a small number of cases deaths may be incorrectly recorded or the notice may not be accurate. In August 2003, the BIRLS Death File was updated to remove erroneous death information which, in the future, will be removed from the BIRLS Death File every January.¹² The BIRLS Death File extract used in this study was pulled on January 9, 2004 and contained deaths through December 19, 2003.

* Sandie Harms-Taylor, Data Management, Compensation and Pension Systems, VA Central Office, telephone conversation with author, March 27, 2003.

The **Medical SAS Inpatient Datasets** (MSID) contain deaths occurring in VHA hospitals.¹³ As indicated above, these deaths are also reported to the VBA to update BIRLS. Because this process requires manual intervention and in some cases validation by the VBA, some deaths occurring in VHA hospitals are not recorded in BIRLS. In our comparison of death dates, we found instances where a death recorded in the MSID was not recorded in the BIRLS Death File. The MSID for fiscal years 1997-2004 were used to extract dates of death through October 20, 2003.

Initially, we used the RMTPRD.MED.SAS.UNIQSSN.BYSTA3N.FYyy.DATA (yy – fiscal year) file to obtain the MSID dates of death. This file is commonly called the Unique Users File and is created by the VISN Support Service Center (VSSC). Because the file also contains dates of death from the BIRLS Death File and the source of the date of death (MSID or BIRLS-DF) cannot be determined from the contents of the Unique Users Files, it was necessary to pull the MSID dates directly from the MSID. We did compare the dates of death in the Unique Users Files to those found in our BIRLS Death File and MSID extracts, and identified dates in the Unique Users Files that were not in either of these extracts. These deaths were included in our study and identified as MSID dates. Some of these dates were included in our samples sent to the NDI and proved to be very unreliable. We suspect that these dates may have been misreported dates of death that are now annually removed from the BIRLS Death File. Thus, these dates were removed from the analysis of the sample results.

The **Medicare Vital Status File** contains demographics, including the date of death, for all veterans for whom the VA receives Medicare data from CMS through the VA Medicare Data Merge Initiative. These are veterans who have even been enrolled in Medicare. The Medicare Vital Status File is received annually. The primary source for dates of death in the Vital Status File is the Social Security Administration, but CMS also receives dates of death from the Railroad Board, from the Medicare claims data, and from death date updates processed by the SSA District Offices.* Even though the SSA is the main source for dates of death in the Medicare Vital Status File, Medicare does contain deaths not identified by the SSA and vice versa. The Medicare Vital Status File used in this study was received from CMS in February 2004 and contained dates of death through December 31, 2003.

The **SSA Death Master File** (SSA DMF) is produced by the Social Security Administration, contains over 70 million deaths, and is updated monthly. The SSA obtains death reports from family members, funeral homes, data exchanges with States and Federal Agencies, postal authorities and financial institutions.¹⁴ The VA maintains a copy of the SSA DMF. We extracted deaths from the VA copy available at the Austin Automation Center (AAC) on January 9, 2004. The extract contained dates of death through December 30, 2003.

The dates of death from all four sources were compared. There were 819,021 veterans in the study cohort with at least one death date in the years 1999 through 2002 and no death date after 2002; 320,417 veterans with death dates only before 1999; and 238,963 veterans with at least one source with a death date after 2002. Because there is approximately an 18-month lag for the release of a calendar year's deaths by the NDI, we only studied the death dates through 2002.

* Research Data Assistance Center (www.resdac.umn.edu), electronic communication with author via the Request Response and Transmission System, May 21, 2003.

Of the 819,021 veterans with a death date between 1999 and 2002, the majority had the same date of death recorded in more than one source: 721,168 (88.1%) had the same date of death in the multiple sources and 61,675 (7.5%) had different dates of death in the multiple sources. Only 36,178 (4.4%) of these veterans had a death date recorded in only one source. The following section describes how samples were drawn from the group of these 819,021 veterans and those veterans without a date of death in any of the four sources for determining the value of using the NDI.

IV. Selection of Samples for Routine NDI Searches and SSA Epidemiological Vital Status Searches

Random samples were selected from the population of veterans with death dates to verify the validity of those dates through an NDI routine search. An NDI routine search only provides date of death information; it does not provide cause of death. Random samples were selected from the population of veterans without death dates to identify deaths not recorded in any of the four sources available within the VA and to validate our methodology of assigning a presumed living status. These samples were sent to both the NDI and the SSA. The SSA provides an Epidemiological Vital Status Search that indicates if the owner of an SSN is dead, presumed living or has an unknown status. The search also validates the social security number and the name and date of birth connected to the social security number. The presumed living status is based on information the SSA collects in their Master Beneficiary File (contains individuals receiving retirement, survivors or disability benefits), Master Earnings File (contains individuals' tax return information), Master Files of Social Security Number Holders and Applications (contains individuals to whom SSA has assigned a Social Security Number) and Supplemental Security Income File (contains individuals receiving Supplemental Security Income).^{*15,16} The cost of the search was \$0.19 per record.

Those veterans without a date of death were grouped into two categories: presumed living and unknown status. For these populations our extracts of the BIRLS Death File, MSID and SSA Death Master File were updated in August 2004, and death dates after 2002 were used to exclude veterans from the presumed living and unknown status categories. There were 374,220 deaths occurring after 2002. A veteran was considered presumed living as of December 31, 2002 if he/she had no date of death in the updated extracts from the four sources of death dates and:

1. was ever enrolled in Medicare (i.e., Medicare Vital Status Record for the veteran generated in December 2003).
2. used VHA health care after December 31, 2002, or
3. was receiving compensation or pension benefits as of December 31, 2003.

Otherwise, the veteran was assigned an unknown status. Table 2 contains a list of the random samples drawn for the NDI and SSA searches.

* Patricia McFadden, Epidemiology Coordinator, Social Security Administration, phone conversation with author, January 26, 2005.

Table 2: Random Samples for NDI and SSA Searches

| Description of Sample | Total Population | Sample Size |
|---|-------------------------|--------------------|
| 1. Multiple sources with the same date of death | 721,168 | 2,376 |
| 2. Multiple sources with differing dates of death. | | |
| a. Differing month or day of death. | 58,908 | 2,267 |
| b. Differing year of death | 2,767 | 500 |
| 3. Only one source with a date of death | 36,178 | 2,154 |
| Total with a death date between 1999-2002 | 819,021 | 7,297 |
| 4. Presumed living | 6,030,458 | 2,000 |
| 5. Unknown status | | |
| a. Under 65 | 1,033,093 | 2,000 |
| b. 65 and older | 27,374 | 2,000 |
| Total without a date of death | 7,090,925 | 6,000 |
| Deaths prior to 1999 | 320,417 | |
| Deaths after 2002* | 374,220 | |
| Total Population | 8,604,583 | 13,297 |
| 6. Activity after Death | 150,294 | 2,000 |
| 7. 'Overall Sample' - Veterans with no date of death before 1999. | 8,284,166 | 3,000 |

*These are the deaths identified after the update to the death sources in August 2004.

Only the presumed living and unknown status samples were sent to the SSA. Some of the populations above were further stratified prior to drawing the random sample. A description of these stratifications is contained in the following sections, which provide the results of the searches.

An NDI search requires submission of the following minimum information: first name, last name, and SSN or date of birth. We provided last name, first name, middle name, date of birth, date of death, sex, and state of residence. An SSA search requires submission of the following minimum information: social security number, last name, first name, and date of birth. We provided last name, first name, middle initial, date of birth and sex. We used several sources to obtain the name and state of residence for the veterans in our samples sent to the NDI and SSA. These included the National Patient Care Data Base (NPCD), Health Eligibility Center (HEC), the Compensation and Pension Master File, and the VBA. Multiple sources were necessary because no one source contained names for all SSNs in the samples. The name attached to an SSN was not the same across all these sources. A manual review of the differences revealed that some were due to different spellings and others appeared to be different individuals. Using the name provided on the Medicare Vital Status File as a "gold standard", we ranked the sources based on their discrepancy rate with the first three characters of the Medicare name. The Compensation and Pension Master File had the lowest rate (1.18%), followed by HEC (2.27%) and the NPCD (2.51%). The VBA had the highest discrepancy rate (5.25%). We used the name from the source with the lowest discrepancy rate for the NDI and SSA search requests.

The SSA search returns only one status for an SSN. If the name or date of birth provided for an SSN does not match the SSA records, no status is provided for the SSN. The NDI search may

return multiple records as possible matches for each record submitted. The NDI uses nine different matching criteria to select possible matches some of which do not require an SSN match.¹¹ Thus, an NDI user must develop a set of criteria to determine if any of the NDI possible matches are true matches. We used the following criteria on the NDI highest-ranking possible match (record that had the greatest number of data items in agreement) to identify a true match:

1. SSN, sex, and two parts of the date of birth match;
2. At least 7 digits of the SSN, date of birth, last name, first name, middle initial if provided by both sources and sex match; or
3. At least 7 digits of the SSN, date of death, last name, first name, middle initial if provided by both sources and sex match. (Only used when a date of death was provided to the NDI).

Criteria 1 is based on the match criteria used by Fleming, et al to merge VA health care administrative databases with Medicare data.¹⁷ The last two criteria were used to account for transposition or one-digit errors in the recording of the social security number. We found no true matches in the other possible match records when the highest-ranking possible match record was not a true match.

V. Analysis of the NDI and SSA Search Results

This section describes the results of NDI and SSA searches for the random samples listed in Table 2 above.

Multiple Sources with the Same Date of Death (Table 2, Row 1)

Table 3 presents the results of NDI searches for the random samples selected from the group of veterans who had a date of death in two or more sources (BIRLS-DF, MSID, Medicare, and SSA) and the dates of death all matched. This group represents 88.1% of all veterans in the study population with a date of death in any source from 1999 through 2002. A small percentage of records in the sample (0.2%) were rejected by the NDI because the minimum set of data elements was not provided on the record (first name, last name, and SSN or first name, last name, and month and year of birth).

There were 12 cases (0.5% of this sample) where the NDI provided no possible matches. Four of these cases were found only in the MSID and BIRLS Death File. Three of the four were over 65, and one would reasonably expect that Medicare would have a record of their enrollment, as would the Social Security Administration. Because we received no record of enrollment from Medicare, these individuals may have an invalid social security number that prevented linking with Medicare and SSA files. The other 8 cases were found in an internal source (MSID and/or BIRLS-DF) and an external source (SSA and/or Medicare). The date of birth and SSN and gender matched across the sources. Assuming that these sources have accurately recorded the death, these cases indicate that the death may not have been recorded properly in the NDI or the death may have occurred outside the U.S. The NDI only has records of deaths occurring within the 50 states, the District of Columbia, Puerto Rico, and the Virgin Islands.¹¹

There were 23 cases (1.0% of this sample) where the NDI returned possible match records, but these records did not meet our criteria for a match noted above. In 14 of these cases, the social security number was found in an external source (Medicare or SSA), but the date of birth and/or sex from the VA sources did not match the demographics from the external source. The highest-ranking NDI possible match record matched on all nine digits of the SSN and the date of death in twelve of these 14 cases. This suggests that the social security number or the demographics were not recorded correctly in one source. Three of these 23 cases were found only in internal sources (BIRLS-DF and MSID). The remaining six of the 23 cases all had dates of death in an external source and some also had a death date in an internal source. The date of birth and sex matched across the internal and external sources. It is possible these dates of death were not in the NDI because they occurred outside the US or were not properly recorded in the NDI.

For the vast majority of veterans in this stratum, a date of death was found through the NDI search (98.3%), and the death date matched the NDI date of death exactly (99.1%) or was within 1-day of the NDI date of death (99.9%). Thus, if a veteran has the same date of death in two or more sources, this death date should be considered accurate. It would not be cost effective to use the NDI to verify these dates.

Multiple Sources with Differing Dates of Death (Table 2, Row 2)

Table 4 provides detail on the stratified random sampling of this category and the NDI search results. These veterans were divided into two groups: veterans where the month or day of death differed across the sources but the year was the same and veterans where the year of death differed across the sources.

Differing month or day of death (Table 2, Row 2a.)

The group of veterans with differing day or month of death but not year of death accounted for 7.2% of all veterans with a date of death recorded in any source from 1999 through 2002. The results of the NDI search for this group were similar to the results for the group of veterans with matching dates of death in multiple sources. Only one record was rejected by the NDI because the minimum set of data elements was not provided. Of the 17 cases with no NDI possible match returned, 15 had matching demographics across the sources, suggesting that the NDI may not have the death recorded or inaccurate demographics. Of the 51 cases with no NDI true match, 30 cases matched the NDI on the full social security number and had at least one sourced date of death within 30 days of the NDI date of death. For 22 of these 30 cases the VA demographics did not match the SSA and Medicare demographics suggesting that the VA demographics might have been incorrect. The remaining 21 of the 51 cases with no NDI true match had VA demographics that matched the demographics on an external source, but the SSN on the highest-ranking NDI possible match record had fewer than 7 digits matching the NDI SSN, suggesting that the NDI had an invalid SSN or was missing the record of death.

For most of the veterans where the month or day of death differed across the sources an NDI true match was found (95.4% to 98.0% depending on the subgroup). Table 5 provides results of

comparing the dates of death for these cases by death date source to the NDI date of death. The MSID dates of death have the highest rate of agreement with the NDI dates of death, followed by BIRLS-DF, Medicare and the SSA. It is not surprising that the MSID has the most accurate date of death, because these are deaths recorded in an inpatient setting. All the sources have similar high rates of agreement when comparing only the month and year of death. It appears that the low rates of agreement on the day of death seen in Medicare and the SSA are due to setting the day of death with either the first, middle or last day of the month when the exact day of death was unknown. We reviewed a distribution of the day of death by source. The distribution of the SSA day of death showed a much larger number of deaths on the 1st and 15th of the month. The distribution of the Medicare day of death showed a much larger number of deaths on the 28th, 30th and 31st of the month. Based on these statistics, a routine was developed to select the date that would most likely match the NDI date:

1. If all sources had a different date:
 - a. If MSID had a date, use the MSID date; otherwise
 - b. If Medicare had a date and the day was not equal to 28, 30 or 31, use the Medicare date; otherwise
 - c. If SSA had a date and the day was not equal to 00, 01 or 15, use the SSA date; otherwise
 - d. If BIRLS Death File had a date use the BIRLS date; otherwise
 - e. If Medicare had a date use the Medicare date; otherwise
 - f. Use the SSA date.
2. If all four sources had a date and the external dates matched and the internal dates matched, but the external dates did not match the internal dates, use the MSID date.
3. If the external sources only had matching dates:
 - a. If the MSID had a date, use the MSID date; otherwise
 - b. Use the external date.
4. If the internal sources only had matching dates use the internal source.
5. If an external source matches an internal source:
 - a. Use the MSID date if the external source matches the MSID date; otherwise
 - b. Use the BIRLS Death File date.

Although the MSID date shows the highest rate of agreement with the NDI date, only 2.9% of veterans in this group have a MSID date of death (1,697 out of 58,908). Therefore, a routine is necessary to select a date from the other sources when there is no MSID date of death. The routine we used significantly improved the agreement rate when compared to the agreement rates on the other three sources. Use of the routine obtains an average sample agreement rate of 94.7% on a date that is within 2 days of the NDI date, and a 99.5% agreement rate on the month and year of death. Thus, it would not be cost effective to use the NDI to verify these dates.

Differing year of death (Table, 2 Row 2b.)

The group of veterans with death dates that differed by year of death was a very small proportion of all deaths (0.3%, 2,767 deaths). This group had a greater NDI reject rate (2.2%), a greater percentage of no possible matches returned from the NDI search (3.0%), and a greater percentage of no true NDI matches (10.5%), than the veterans where the year of death did not differ. (See Table 4) All of the NDI rejects had VA demographics that did not match SSA and

Medicare demographics; 11 of the 15 (73.3%) veterans with no possible matches had VA demographics that did not match either the SSA or Medicare demographics; and 42 of the 52 (80.8%) veterans with no true NDI match had VA demographics that did not match either SSA or Medicare demographics. Veterans with true matches only had 71 out of 418 (17.0%) with VA demographics that did not match either SSA or Medicare demographics. Thus, the higher percentage of rejects and unmatched NDI death dates is most likely due to invalid demographics or SSN on file for the veteran. VA demographics were used for the NDI search.

Table 6 contains the NDI date of death agreement rates by source for veterans with differing year of death across the sources. The MSID dates have the best match rate (89.5%), followed by Medicare (77.2%), SSA (65.2%) and then BIRLS-DF (26.2%). We developed a routine that improves the overall match rate significantly:

1. If the BIRLS-DF date of death equals the Medicare or SSA date of death and the either the BIRLS-DF, Medicare or SSA demographics match the VA demographics, use the BIRLS-DF date of death; otherwise
2. If the MSID demographics match the VA demographics, use the MSID date of death; otherwise
3. If the Medicare demographics match the VA demographics, use the Medicare date of death; otherwise
4. If the SSA demographics match the VA demographics use the SSA date of death; otherwise
5. If the BIRLS-DF demographics match the VA demographics use the BIRLS-DF date of death; otherwise
6. Use the MSID date of death if there is a MSID date of death; otherwise
7. Use the Medicare date of death if there is a Medicare date of death; otherwise
8. Use the SSA date of death if there is an SSA date of death; otherwise
9. Use the BIRLS-DF date of death if there is a BIRLS-DF date of death.

This routine selected date that had an 87.3% match rate with the NDI date of death; 87.8% match rate within two days of the NDI date of death (367 out of 418); and a 90.9% match rate with the NDI year and month of death. Only 2 out of the 418 veterans with a true NDI match had no date of death on any source that at least matches the NDI year and month of death. The most significant issue with this group was the low yield rate of an NDI true match (84.3%). Because using a routine provides a good mechanism to select the most likely date of death and this group of veterans is a very small percentage of all deceased veterans, use of the NDI to validate dates of death for this group would not significantly improve the accuracy of death ascertainment in the larger population.

Only One Source with a Date of Death (Table 2, Row 3)

There were 36,178 veterans with a date of death in only one source (4.4% of all deaths). Table 7 contains the NDI search results for these veterans. Medicare and SSA provided 79.3% of these deaths. BIRLS-DF and MSID showed a significantly greater rate of veterans with no NDI possible matches or no NDI true matches. Expanding the criteria of a true match to include a match on name and date of birth increased the BIRLS-DF rate of true matches to 78.8% from

74.1% and the MSID to 83.3% from 82.2% with little effect on the agreement rates of source death dates with the NDI date of death. These rates were still lower than the Medicare and SSA true match rates. For those records with an NDI true match, the date of death agreement rates are comparable across the different sources (92.5% to 97.3% for an exact match and 95.6% to 98.6% for a match within two days of the NDI date), and have an average sample rate slightly better than the rate for veterans with differing dates in multiple sources when a routine is used to pick the most likely date of death (96.2% versus 94.7%). Due to the small size of this category and the good match rates for the largest subgroups (Medicare and SSA), we see little benefit to validating these dates with an NDI search on an ongoing basis.

In our analysis of the distribution by year and month of the date of deaths by source, we identified several months in which the SSA had an unusually low number of deaths. In these same months Medicare had an unusually high number of deaths found only in Medicare. We suspect that there was a problem in the monthly process to update the SSA Death Master File either at the SSA or the VA. The issue has been reported to National Data Systems, but a cause has yet to be identified. In all systems there is the possibility of error; therefore, use of multiple sources reduces the chance to introduce bias in death ascertainment caused by those errors.

Presumed Living (Table 2, Row 4)

This category includes veterans who were presumed alive on December 31, 2002, i.e. they either received compensation or pension benefits after that date; had VHA activity after that date; or where enrolled in Medicare. It is the largest category in our cohort (6,030,458 veterans, 70.1%). The random sample from this category as expected had the greatest rates of records with no NDI possible matches (32.2%) and no NDI true matches (67.8%). The NDI search identified no deaths in the sample from this category. The SSA Epidemiological Vital Status Search identified 90.6% of these veterans as living, 7.2% of these as unknown status, 0.6% as deceased (all with death dates in 2004), and the remaining 1.6% with invalid demographics (name does not match the name associated with the SSN at the SSA or the SSN is invalid). Thus, our method of categorizing a veteran as presumed living is quite accurate, and there is no benefit of performing an NDI search to identify deaths in this category.

Unknown Status (Table 2, Row 5)

The unknown status category includes all veterans without a date of death in any source through mid-2004 and not categorized as presumed living. These veterans represent 12.3% of the cohort. The majority were under 65 as of January 1, 1999 (94.7%). The NDI and SSA search results for the veterans under 65 versus veterans 65 and older group were significantly different. Tables 8 and 9 contain the results. The under-65 group looked very similar to the presumed living veterans. Only 6 deaths (0.3%) were identified in this group as deceased from the NDI search. The SSA results identified 87.6% of this group as presumed living and 8.0% with an unknown status. Only 0.4% was identified as deceased. An NDI search or an SSA search for this group would cost over \$150,000 per year and would yield a miniscule number of deaths. Thus, use of the NDI to ascertain additional deaths in the under-65 group would not be cost effective.

The veterans who were 65 or older in this category were significantly different from those under 65. The SSA search identified only 5.3% of the 65 or older veterans as presumed living. More importantly, 74.0% had invalid demographics (i.e., VA name did not match the name associated with the SSN at the SSA or the SSN was invalid). Because of this, we expanded the NDI true match criteria to include records if the first name, last name, middle initial and date of birth matched the NDI information for this group and also for the group under 65 reported above. This produced a true match rate of 9.0% or 179 deaths. The SSA search confirmed 6 of the 179 deaths, categorized 12 as unknown status, and reported that the remaining 161 had invalid demographics or an invalid SSN. The high rate of invalid SSN or demographics most likely is the reason we do not have Medicare data for these veterans, and leaves the NDI as the most viable source for death ascertainment for this group. The SSA search identified 3 additional deaths in this group prior to 2003. The cost of searching the NDI for this group would be approximately \$21,350 for the years 1999 – 2003 (28,000 records * \$0.15 per record per year * 5 years + \$350 flat fee). The NDI search costs for each following year would be approximately \$4,550. This does not include personnel costs and equipment costs for annually identifying this group and building a database of the NDI search results that would be available to researchers. These costs would at least double the NDI search costs. Because this group is such a small percentage of the veteran population (0.3%), using the NDI on an ongoing basis to identify deaths in this group would not be cost effective.

Activity After Death (Table 2, Row 6)

We looked at veterans who had inpatient or outpatient activity after death either in Medicare or the VHA to determine if activity after death could be used to identify incorrectly recorded deaths. The Medicare data were only available through calendar year 2001 at the time of our study. We ascertained VHA activity through mid-2004 using the Unique Users Files described previously. There were 150,294 veterans with activity after the most recent recorded date of death. Table 10 provides the NDI search results for these veterans. The percentage of these deaths found in the NDI is 96% or greater for each subgroup in the sample except for the subgroup with activity more than a year after death. Only 86.6% of the deaths in the subgroup were found in the NDI. We reviewed those deaths for which the NDI returned no possible matches or true matches. A much larger percentage of these deaths were found in only one source (38.2%) than for those deaths with an NDI true match (3.3%). In the subgroup with activity more than a year after death, 61.2% of those veterans without an NDI true match had a date of death in only one source. Of the 89 veterans in the full sample without an NDI true match whose records were not rejected by the NDI, 34 had only one source with a date of death, and for 26 of these 34 (76.4%) BIRLS-DF or the MSID was the single source. This supports the previous findings in the results from the sample of deaths in only one source. When BIRLS-DF or the MSID was the single source for a date of death, it was less likely that the death would be found in the NDI when compared to single-sourced dates of death from Medicare or the SSA. Thus, deaths found only in BIRLS-DF or the MSID with activity more than a year after death are probably recorded inaccurately or the veteran's SSN or demographics are incorrect.

The rates of agreement between the death dates found in the VA sources and the NDI death date are reported in Table 11. These rates were high (92.4% - 97.7%) except for the category of activity after death from 1 to 31 days after the date of death in Medicare (89.0%) and the SSA

(76.6%). All sources had high rates of agreement when only matching on year and month of death (97.4% - 100%). We suspect that Medicare and the SSA dates of death show a lower rate of agreement on the date of death in the 1 to 31 days category because they are more likely to “plug” the day of death if the exact date is not reported to them. We did review the distribution of day of death for these cases. There was a greater number of deaths recorded on the 1st and 15th of the month in the SSA for deaths in this category than on other days. We did not see this same distribution in the category of deaths with activity 32 to 60 days after the date of death.

Based on our analysis, we did not find that activity after death signified a significantly greater likelihood that the record of death would not be found through an NDI search, except in those instances where the activity occurred more than a year after death and the death was found in only BIRLS-DF or the MSID.

Overall Sample (Table 2, Row 7)

Lastly we took a random sample of 3,000 veterans out of the 8,284,166 veterans who had no date of death or had at least one date of death after December 31, 1998 to calculate the sensitivity of using the four sources for death dates available within the VA as compared to using an NDI search. The NDI rejected 2 (0.1%) of the sample records and returned no possible matches for 1,808 (60.3%) of the sample records. Of the veterans with possible match records returned, 292 (9.7% of the sample) had true matches (i.e., identified as deceased) and 898 (30.0% of the sample) had no true match. Table 12 details how this sample was categorized within the VA as presumed living, unknown status or deceased using the criteria discussed above, and the number of NDI identified deaths in each of these categories. There was the expected number of NDI deaths found in the unknown status category (2 deaths, 0.4%). There were 3 NDI deaths identified in the presumed living category. We would expect to find no deaths in this category, but the number found was a very small percent of the category, 0.1%. In the first case the individual was under 65 and had VHA activity almost two years after the date of death recorded in the NDI. The other two cases appear to be an issue with demographics. VA demographics were used for the NDI search. The demographics for the SSN in Medicare did not match the VA demographics for the same SSN.

Use of the four sources to ascertain death (BIRLS-DF, MSID, SSA, and Medicare) identified 98.3% of all NDI deaths. Five deaths recorded in these sources were not found through the NDI search using our true match criteria. One of these records had a date of death in SSA, Medicare, and BIRLS-DF and had matching demographics across these sources, but there were no NDI possible match records returned. We classified this as a death not recorded in the NDI. A second record also had a date of death in SSA, Medicare, and BIRLS-DF and matching demographics across these sources. The NDI returned possible match records, but none met our true match criteria. A review of these records identified a record that was most likely the death record in the NDI, but the record had one unmatched digit on the SSN and the middle initial did not match, thus it did not meet our criteria for a true match. The third record had only a date of death in the SSA, but the SSA date of birth (1955) did not match the VA date of birth (1910). The NDI did return possible match records and the highest-ranking record had the same date of death as recorded in the SSA. The fourth record had only a date of death in Medicare, but the Medicare gender (female) and date of birth (1927) did not match the VA gender and date of birth

(1926). The NDI did return possible match records and the highest-ranking record had a date of death within 23-days of the date recorded in Medicare. The date recorded in Medicare was the last day of the month, so it may have been a “plugged” day. The fifth record had only a date of death recorded in BIRLS-DF, which matched the date of death on the highest-ranking NDI possible match record. We did not pick this record up as a true match because the VA SSN matched on only 2-digits of the NDI SSN. The VA full name, gender, and date of birth matched these same data elements on the NDI. The Medicare gender (female) and date of birth (1926) did not match the VA gender and date of birth (1922). Thus, four of these deaths were recorded in the NDI, but our true match criteria prevented an automated identification of these deaths due to possibly incorrect recording of the SSN or demographics.

Using the NDI as the gold standard and not adjusting for the cases detailed above, we calculated the sensitivity and specificity of using the four sources available within the VA to identify deaths (Table 13). Sensitivity measures the percentage of veterans classified as deceased according to the NDI that are identified as deceased in the VA sources. Specificity measures the percentage of veterans who are not identified in the NDI as deceased that are not deceased according to the VA sources. The sensitivity and specificity were very high, 98.3% and 99.8% respectively. Additionally, when comparing the actual dates of death in the four sources to the date of death in the NDI on true matches, there was a very high rate of agreement (Table 14). We reported in Table 14 results of using the routine detailed in Section V that identifies the most likely date of death when the date of death differed among the sources for a veteran. The routine performed extremely well. The date match rates across the sources ranged from 95.9% to 100%. The rates for a date match within two days ranged from 97.9% to 100%; and for a month and year match from 99.6% to 100%. Thus, the NDI search results for this random sample of the whole population confirmed the findings from all the previous samples. Use of the four sources (BIRLS-DF, MSID, Medicare, SSA) was basically equivalent to use of an NDI search for death ascertainment in this sample, further supporting our conclusion that it would not be cost effective to use the NDI in building a death registry.

Summary of Our Analyses

Based on our analysis of the NDI and SSA search results, we conclude that it would be feasible to build a Death Registry using the Medicare Vital Status File, the SSA Death Master File, the BIRLS Death File, and the MSID. It would not be cost effective to augment this Registry with NDI searches on an ongoing basis. Using VHA and Medicare utilization dates and dates of compensation and pension benefits payments, a presumed living status could be established. Adding this information to the Registry would provide useful information in assessing the vital status of veterans. Researchers would be able to classify their cohort into one of three categories: presumed living, unknown status, and deceased. Using the results of our analyses, researchers could select the most likely date of death when dates of death do not match across the sources and develop an assessment of the completeness and accuracy of death ascertainment for their cohort. Based on this assessment, researchers could determine whether use of an NDI search would improve the quality of death ascertainment in their cohort.

Limitations

We provided VA demographics (date of birth and gender) in our search requests to the SSA and the NDI. As noted above the demographics for the same SSN were not consistent across the VA data sources. Thus, we used the most commonly occurring date of birth and gender. These demographics occasionally did not match the demographics from the SSA, Medicare, BIRLS-DF and MSID files. Match results might be improved by sending a record to the SSA and NDI for each unique combination of date of birth and gender for an SSN. Additionally, our match criteria for an NDI true match were conservative. Loosening these criteria may increase the number of NDI true matches, improving rates of agreement between the death sources available within the VA and the NDI.

Table 3: NDI Search Results from Cases with Dates of Death in Multiple Sources That Match

| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O |
|--------------------------|--------------------|-------------------------|---------------------------------------|-------------|---------------------------------------|-------------|--------------------------------------|-------------|-----------------------------------|-------------|-------------------------------------|-------------------|--|-------------------|
| Sources with Death Dates | Number of Veterans | Number in Random Sample | Number of Records Rejected by the NDI | % of Sample | Number of Records with no NDI Matches | % of Sample | Number of Records with no True Match | % of Sample | Number of Records with True Match | % of Sample | True Match with Date of Death Match | % of True Matches | True Match with Date of Death Match within One Day | % of True Matches |
| BIRLS-DF and MSID | 1,454 | 200 | 0 | 0.0% | 4 | 2.0% | 3 | 1.5% | 193 | 96.5% | 193 | 100.0% | 193 | 100.0% |
| SSA and Medicare | 116,155 | 312 | 1 | 0.3% | 2 | 0.6% | 7 | 2.2% | 302 | 96.8% | 294 | 97.4% | 301 | 99.7% |
| MSID and External | 7,304 | 200 | 0 | 0.0% | 0 | 0.0% | 1 | 0.5% | 199 | 99.5% | 196 | 98.5% | 198 | 99.5% |
| BIRLS-DF and External | 501,353 | 1,390 | 4 | 0.3% | 5 | 0.4% | 10 | 0.7% | 1371 | 98.6% | 1,361 | 99.3% | 1,371 | 100.0% |
| BIRLS-DF, MSID, External | 94,902 | 274 | 0 | 0.0% | 1 | 0.4% | 2 | 0.7% | 271 | 98.9% | 271 | 100.0% | 271 | 100.0% |
| Total | 721,168 | 2,376 | 5 | 0.2% | 12 | 0.5% | 23 | 1.0% | 2,336 | 98.3% | 2,315 | 99.1% | 2,334 | 99.9% |

Row Descriptions:

BIRLS-DF and MSID = Veteran had the same death date in BIRLS-DF and MSID. There was no death date found in SSA or Medicare.
 SSA and Medicare = Veteran had the same death date in SSA and Medicare. There was no death date found in BIRLS-DF or MSID.
 MSID and External = Veteran had the same death date in MSID and either Medicare or SSA or both. There was no death date in BIRLS-DF.
 BIRLS-DF and External = Veteran had the same death date in BIRLS-DF and either Medicare or SSA or both. There was no death date in BIRLS-DF.
 BIRLS-DF, MSID and External = Veteran had the same death date in BIRLS-DF, MSID and either Medicare or SSA or both.

Column Descriptions:

- A. See Row Descriptions.
- B. Number of Veterans: The number of veterans with dates of deaths in the sources identified in Column A.
- C. Number in Random Sample: The number of veterans randomly selected from the veterans in Column B for an NDI search. One record was sent to the NDI per veteran.
- D. Number of Records Rejected by the NDI: The number of records rejected by the NDI because the minimum set of data elements was not provided on the record (first name, last name, and SSN or first name, last name, and month and year of birth).
- E. % of Sample: Percent of records in the random sample rejected by the NDI. Column D divided by Column C.
- F. Number of Records with no NDI Matches: Number of records for which the NDI returned no possible matches.
- G. % of Sample: Percent of records in the random sample for which the NDI returned no possible matches. Column F divided by Column C.
- H. Number of Records with no NDI True Match: The number of records for which the NDI returned possible matches, but none were deemed to be true matches. See description of a true match in Section V.
- I. % of Sample: Percent of records in the random sample with no NDI true match. Column H divided by Column C.
- J. Number of Records with True Match: The number of records for which the highest-ranking NDI possible match was deemed to be a true match.
- K. % of Sample: Percent of records in the random sample with an NDI true match. Column J divided by Column C.
- L. True Match with Date of Death Match: The number of true matches where the date of death on the NDI true match record agrees with the date of death in the VA sources.
- M. % of Sample: Percent of records where the VA sourced date of death matches the NDI date of death exactly. Column L divided by Column J.
- N. True Match with Date of Death Match within One Day: The number of true matches where the date of death on the NDI true match record is within one day of the date of death in the VA sources.
- O. % of Sample: Percent of records where the VA sourced date of death is within one day of the NDI date of death. Column N divided by Column J.

Table 4: NDI Search Results from Cases with Dates of Death in Multiple Sources That Differ Across the Sources

| A | B | C | D | E | F | G | H | I | J | K |
|---|--------------------|-------------------------|---------------------------------------|-------------|---------------------------------------|-------------|--------------------------------------|-------------|-----------------------------------|-------------|
| | Number of Veterans | Number in Random Sample | Number of Records Rejected by the NDI | % of Sample | Number of Records with no NDI Matches | % of Sample | Number of Records with no True Match | % of Sample | Number of Records with True Match | % of Sample |
| Death Dates Matching Among Sources | | | | | | | | | | |
| No Death Dates Match | 16,602 | 593 | 0 | 0.0% | 8 | 1.3% | 19 | 3.2% | 566 | 95.4% |
| An External Date Matches an Internal Date | 20,401 | 653 | 1 | 0.2% | 3 | 0.5% | 11 | 1.7% | 638 | 97.7% |
| Only External Match | 17,976 | 621 | 0 | 0.0% | 5 | 0.8% | 14 | 2.3% | 602 | 96.9% |
| Only Internal Match | 2,185 | 200 | 0 | 0.0% | 0 | 0.0% | 4 | 2.0% | 196 | 98.0% |
| Internal Dates Match and External Dates | 1,744 | 200 | 0 | 0.0% | 1 | 0.5% | 3 | 1.5% | 196 | 98.0% |
| Total | 58,908 | 2,267 | 1 | 0.0% | 17 | 0.7% | 51 | 2.2% | 2,198 | 97.0% |
| Year of Death Differs | 2,767 | 496 | 11 | 2.2% | 15 | 3.0% | 52 | 10.5% | 418 | 84.3% |

Row Descriptions:

No Death Dates Match: The dates of death found in the sources are all different dates.

An External Date Matches an Internal Date: At least one external date of death (SSA or Medicare date) matches an internal date of death (BIRLS-DF or MSID date of death).

Only External Match: The SSA date of death matches the Medicare date of death, but this date does not match either the BIRLS-DF or MSID date of death and the BIRLS-DF date of death does not match the MSID date of death.

Only Internal Match: The BIRLS-DF date of death matches the MSID date of death, but this date does not match either the SSA or Medicare date of death and the SSA date of death does not match the Medicare date of death.

Internal Dates Match and External Dates Match: The internal dates of death match and the external dates of death match, but the internal date of death does not match the external date of death.

Year of Death Differs: The year of death among the sources with a date of death is not the same year.

Column Descriptions:

A. See Row Descriptions.

B. Number of Veterans: The number of veterans in the categories identified in Column A.

C. Number in Random Sample: The number of veterans randomly selected from the veterans in Column B for an NDI search. One record was sent to the NDI per veteran.

D. Number of Records Rejected by the NDI: The number of records rejected by the NDI because the minimum set of data elements was not provided on the record (first name, last name, and SSN or first name, last name, and month and year of birth).

E. % of Sample: Percent of records in the random sample rejected by the NDI. Column D divided by Column C.

F. Number of Records with no NDI Matches: Number of records for which the NDI returned no possible matches.

G. % of Sample: Percent of records in the random sample for which the NDI returned no possible matches. Column F divided by Column C.

H. Number of Records with no NDI True Match: The number of records for which the NDI returned possible matches, but none were deemed to be true matches. See description of a true match in Section V.

I. % of Sample: Percent of records in the random sample with no NDI true match. Column H divided by Column C.

J. Number of Records with True Match: The number of records for which the NDI returned possible matches and the highest-ranking possible match was deemed to be a true match.

K. % of Sample: Percent of records in the random sample with an NDI true match. Column J divided by Column C.

Table 5: Agreement Rates with NDI Date of Death for Cases with Dates of Death in Multiple Sources That Differ in Month or Day across the Sources

| | Source of Date of Death | | | | |
|--|-------------------------|----------|----------|--------|---------|
| | SSA | Medicare | BIRLS-DF | MSID | Routine |
| Number of Deaths in Source | 2,083 | 1,914 | 2,014 | 573 | |
| % With Death Date Matching the NDI Death Date | | | | | |
| No Death Dates Match | 35.3% | 33.2% | 58.5% | 81.8% | 77.9% |
| An External Matches an Internal Date | 48.4% | 59.8% | 87.5% | 88.3% | 95.3% |
| Only External Match | 71.1% | 71.1% | 26.4% | 100.0% | 72.4% |
| Only Internal Match | 1.2% | 0.8% | 98.0% | 98.0% | 98.0% |
| Internal Dates Match and External Dates Match | 3.1% | 3.1% | 96.9% | 96.9% | 96.9% |
| Average for the Sample | 43.9% | 48.6% | 65.7% | 94.6% | 84.9% |
| % With Death Date Within 2 Days of the NDI Death Date | | | | | |
| No Death Dates Match | 53.9% | 43.5% | 82.8% | 90.9% | 89.9% |
| An External Matches an Internal Date | 68.4% | 68.4% | 100.0% | 100.0% | 98.7% |
| Only External Match | 91.7% | 91.7% | 64.7% | 100.0% | 91.9% |
| Only Internal Match | 24.5% | 15.6% | 99.0% | 99.0% | 99.0% |
| Internal Dates Match and External Dates Match | 64.7% | 66.8% | 95.3% | 98.8% | 100.0% |
| Average for the Sample | 67.2% | 67.0% | 84.6% | 99.1% | 94.7% |
| % With Year and Month of Death Matching the NDI Year and Month of Death | | | | | |
| No Death Dates Match | 98.6% | 98.9% | 94.9% | 100.0% | 99.3% |
| An External Matches an Internal Date | 95.4% | 95.4% | 100.0% | 100.0% | 100.0% |
| Only External Match | 98.8% | 98.8% | 92.9% | 100.0% | 99.0% |
| Only Internal Match | 94.5% | 91.4% | 99.5% | 99.5% | 99.5% |
| Internal Dates Match and External Dates Match | 99.4% | 99.8% | 98.7% | 98.8% | 100.0% |
| Average for the Sample | 98.3% | 98.3% | 96.5% | 99.5% | 99.5% |

Row Descriptions:

Number of Deaths in Source: The number of death dates found in the source for the group of veterans with death dates in multiple sources that differ.

No Death Dates Match: The dates of death found in the sources are all different dates.

An External Date Matches an Internal Date: At least one external date of death (SSA or Medicare date of death) matches an internal date of death (BIRLS-DF or MSID date of death).

Only External Match: The SSA date of death matches the Medicare date of death, but this date does not match either the BIRLS-DF or MSID date of death and the BIRLS-DF date of death does not match the MSID date of death.

Only Internal Match: The BIRLS-DF date of death matches the MSID date of death, but this date does not match either the SSA or Medicare date of death and the SSA date of death does not match the Medicare date of death.

Internal Dates Match and External Dates Match: The internal dates of death match and the external dates of death match, but the internal date of death does not match the external date of death.

Average for the Sample: This average is a raw average; it is not weighted by probability of selection.

Column Descriptions:

SSA: Source of date is the SSA Death Master File.

Medicare: Source of date is the Medicare Vital Status File.

BIRLS-DF: Source of date is the BIRLS Death File.

MSID: Source of date is the MSID.

Routine: Date is selected from one of the four sources (SSA, Medicare, BIRLS-DF, MSID) based on an automated routine.

Table 6: Agreement Rates with NDI Date of Death for Cases with Dates of Death in Multiple Sources That Differ in Year across the Sources

| Source | % Dates Agree | % Month and Year Agree | Number of Death Dates | % Matching VA Demographics |
|----------|---------------|------------------------|-----------------------|----------------------------|
| Medicare | 77.2% | 81.9% | 337 | 86.65% |
| SSA | 65.2% | 69.0% | 394 | 83.76% |
| BIRLS-DF | 26.2% | 28.2% | 390 | 73.59% |
| MSID | 89.5% | 90.7% | 86 | 98.84% |
| Routine | 87.3% | 90.9% | 418 | N/A |

Row Descriptions:

SSA: Source of death date is the SSA Death Master File.

Medicare: Source of death date is the Medicare Vital Status File.

BIRLS-DF: Source of death date is the BIRLS Death File.

MSID: Source of death date is the MSID.

Routine: Date is selected from one of the four sources (SSA, Medicare, BIRLS-DF, MSID) based on an automated routine.

Column Descriptions:

% Dates Agree: Percent of death dates found in the Medicare Vital Status File that agrees with the NDI date.

% Month and Year Agree: Percent of death dates found in the Medicare Vital Status File that agree with the NDI date on the month and year of death.

Number of Death Dates: Number of death dates found in the source for veterans with dates of death in multiple sources that differ.

% Matching VA Demographics: Percent of records where the demographics on the source match the demographics obtained in generation of the study population.

Table 7: NDI Search Results from Cases with a Date of Death Only in One Source

| A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q |
|----------|--------------------|-------------------------|---------------------------------------|-------------|---------------------------------------|-------------|--------------------------------------|-------------|-----------------------------------|-------------|-------------------------------------|-------------------|---|-------------------|---|-------------------|
| Source | Number of Veterans | Number in Random Sample | Number of Records Rejected by the NDI | % of Sample | Number of Records with no NDI Matches | % of Sample | Number of Records with no True Match | % of Sample | Number of Records with True Match | % of Sample | True Match with Date of Death Match | % of True Matches | True Match with Date of Death Match within Two Days | % of True Matches | True Match with Year and Month of Death Match | % of True Matches |
| Medicare | 9,570 | 547 | 2 | 0.4% | 12 | 2.2% | 7 | 1.3% | 526 | 96.5% | 500 | 95.1% | 508 | 96.6% | 524 | 99.6% |
| SSA | 19,117 | 1,048 | 0 | 0.0% | 50 | 4.8% | 12 | 1.1% | 986 | 94.1% | 921 | 93.4% | 943 | 95.6% | 978 | 99.2% |
| BIRLS-DF | 6,656 | 359 | 0 | 0.0% | 55 | 15.3% | 38 | 10.6% | 266 | 74.1% | 246 | 92.5% | 258 | 97.0% | 263 | 98.9% |
| MSID | 835 | 90 | 0 | 0.0% | 11 | 12.2% | 5 | 5.6% | 74 | 82.2% | 72 | 97.3% | 73 | 98.6% | 74 | 100.0% |
| Total | 36,178 | 2,044 | 2 | 0.1% | 128 | 6.3% | 62 | 3.0% | 1,852 | 90.7% | 1,739 | 93.9% | 1,782 | 96.2% | 1,839 | 99.3% |

Row Descriptions:

Medicare: Date of death only found in the Medicare Vital Status File.
 SSA: Date of death only found in the SSA Death Master File.
 BIRLS-DF: Date of death only found in the BIRLS Death File.
 MSID: Date of death only found in the MSID.

Column Descriptions:

- A. See Row Descriptions.
- B. Number of Veterans: The number of veterans in the categories identified in Column A.
- C. Number in Random Sample: The number of veterans randomly selected from the veterans in Column B for an NDI search. One record was sent to the NDI per veteran.
- D. Number of Records Rejected by the NDI: The number of records rejected by the NDI because the minimum set of data elements was not provided on the record.
- E. % of Sample: Percent of records in the random sample rejected by the NDI. Column D divided by Column C.
- F. Number of Records with no NDI Matches: Number of records for which the NDI returned no possible matches.
- G. % of Sample: Percent of records in the random sample for which the NDI returned no possible matches. Column F divided by Column C.
- H. Number of Records with no NDI True Match: The number of records for which the NDI returned possible matches, but none were deemed to be true matches. See description of a true match in Section V.
- I. % of Sample: Percent of records in the random sample with no NDI true match. Column H divided by Column C.
- J. Number of Records with True Match: The number of records for which the NDI returned possible matches and the highest-ranking match was deemed to be a true match.
- K. % of Sample: Percent of records in the random sample with an NDI true match. Column J divided by Column C.
- L. True Match with Date of Death Match: The number of true matches where the date of death on the NDI true match record agrees with the date of death in the VA source.
- M. % of Sample: Percent of records where the VA sourced date of death matches the NDI date of death exactly. Column L divided by Column J.
- N. True Match with Date of Death Match within One Day: The number of true matches where the date of death on the NDI true match record is within two days of the date of death in the VA source.
- O. % of Sample: Percent of records where the VA sourced date of death is within two days of the NDI date of death. Column N divided by Column J.
- P. True Match with Year and Month of Death Match: The number of true matches where the year and month of death on the NDI true match record matches the year and month of death in the VA source.
- Q. % of Sample: Percent of records where the VA sourced year and month of death matches the NDI year and month of death. Column N divided by Column J.

Table 8: NDI Search Results for Veterans with an Unknown Status

| A | B | C | D | E | F | G | H | I | J | K |
|--------------|--------------------|-------------------------|---------------------------------------|-------------|---------------------------------------|-------------|--------------------------------------|-------------|-----------------------------------|-------------|
| Age | Number of Veterans | Number in Random Sample | Number of Records Rejected by the NDI | % of Sample | Number of Records with no NDI Matches | % of Sample | Number of Records with no True Match | % of Sample | Number of Records with True Match | % of Sample |
| 65 and Older | 27,374 | 2,000 | 3 | 0.2% | 1166 | 58.3% | 679 | 34.0% | 179 | 9.0% |
| Under 65 | 1,033,093 | 2,000 | 0 | 0.0% | 1424 | 71.2% | 570 | 28.5% | 6 | 0.3% |

Row Descriptions:

65 and Older: Veterans 65 or older on January 1, 1999.

Under 65: Veterans under 65 on January 1, 1999.

Column Descriptions:

A. See Row Descriptions.

B. Number of Veterans: The number of veterans in the categories identified in Column A.

C. Number in Random Sample: The number of veterans randomly selected from the veterans in Column B for an NDI search. One record was sent to the NDI per veteran.

D. Number of Records Rejected by the NDI: The number of records rejected by the NDI because the minimum set of data elements was not provided on the record (first name, last name, and SSN or first name, last name, and month and year of birth).

E. % of Sample: Percent of records in the random sample rejected by the NDI. Column D divided by Column C.

F. Number of Records with no NDI Matches: Number of records for which the NDI returned no possible matches.

G. % of Sample: Percent of records in the random sample for which the NDI returned no possible matches. Column F divided by Column C.

H. Number of Records with no NDI True Match: The number of records for which the NDI returned possible matches, but none were deemed to be true matches. See description of a true match in Section V.

I. % of Sample: Percent of records in the random sample with no NDI true match. Column H divided by Column C.

J. Number of Records with True Match: The number of records for which the NDI returned possible matches and the highest-ranking possible match was deemed to be a true match.

K. % of Sample: Percent of records in the random sample with an NDI true match. Column J divided by Column C.

Table 9: SSA Search Results for Veterans with an Unknown Status

| SSA Status | 65 and Older | | Under 65 | |
|----------------------------------|--------------|---------|----------|---------|
| | Number | Percent | Number | Percent |
| Presumed Living | 105 | 5.3% | 1,751 | 87.6% |
| Dead | 16 | 0.8% | 7 | 0.4% |
| Unknown Status | 399 | 20.0% | 160 | 8.0% |
| Name Does Not Match SSN | 1,022 | 51.1% | 63 | 3.2% |
| Date of Birth Does Not Match SSN | 121 | 6.1% | 8 | 0.4% |
| Invalid SSN | 293 | 14.7% | 3 | 0.2% |
| Input Name is Blank | 44 | 2.2% | 8 | 0.4% |
| Total Sample | 2,000 | 100.0% | 2,000 | 100.0% |

Row Descriptions:

Presumed Living: Presumed living based on SSA administrative records.

Dead: Person is deceased according to the SSA records.

Unknown Status: SSA has no record of death and is unable to support a presumed living status.

Name Does Not Match SSN: The name provided did not match the name on file at the SSA for the SSN (spelling tolerances are incorporated into the SSA's name matching algorithm).

Date of Birth Does Not Match SSN: Date of birth provided did not match the date of birth on file at the SSA for the SSN.

Invalid SSN: The SSN is invalid or has never been issued.

Input Name is Blank: Name on the record submitted to the SSA is blank.

Column Descriptions:

65 and Older: Number and Percent of veterans 65 or older on 1/1/99 classified with an unknown status out of the total sample.

Under 65: Number and Percent of veterans under 65 on 1/1/99 classified with an unknown status out of the total sample.

Table 10: NDI Results for Veterans with Activity after Death

| A | B | C | D | E | F | G | H | I | J | K |
|----------------------------|--------------------|-------------------------|---------------------------------------|-------------|---------------------------------------|-------------|--------------------------------------|-------------|-----------------------------------|-------------|
| Number of Days After Death | Number of Veterans | Number in Random Sample | Number of Records Rejected by the NDI | % of Sample | Number of Records with no NDI Matches | % of Sample | Number of Records with no True Match | % of Sample | Number of Records with True Match | % of Sample |
| 1-31 | 62,602 | 200 | | 0.0% | 1 | 0.5% | 2 | 1.0% | 197 | 98.5% |
| 32-61 | 18,517 | 374 | | 0.0% | 5 | 1.3% | 4 | 1.1% | 365 | 97.6% |
| 62-180 | 32,622 | 687 | 1 | 0.1% | 10 | 1.5% | 3 | 0.4% | 673 | 98.0% |
| 180-365 | 18,674 | 367 | 1 | 0.3% | 11 | 3.0% | 4 | 1.1% | 353 | 96.2% |
| 366+ | 17,879 | 372 | 1 | 0.3% | 26 | 7.0% | 23 | 6.2% | 322 | 86.6% |
| Total | 150,294 | 2,000 | 3 | 0.2% | 53 | 2.7% | 36 | 1.8% | 1,910 | 95.5% |

Row Descriptions:

Number of days between the most recent date of death in any source and the most recent activity in Medicare or the VHA.

Column Descriptions:

- A. See Row Descriptions.
- B. Number of Veterans: The number of veterans in the categories identified in Column A.
- C. Number in Random Sample: The number of veterans randomly selected from the veterans in Column B for an NDI search. One record was sent to the NDI per veteran.
- D. Number of Records Rejected by the NDI: The number of records rejected by the NDI because the minimum set of data elements was not provided on the record (first name, last name, and SSN or first name, last name, and month and year of birth).
- E. % of Sample: Percent of records in the random sample rejected by the NDI. Column D divided by Column C.
- F. Number of Records with no NDI Matches: Number of records for which the NDI returned no possible matches.
- G. % of Sample: Percent of records in the random sample for which the NDI returned no possible matches. Column F divided by Column C.
- H. Number of Records with no NDI True Match: The number of records for which the NDI returned possible matches, but none were deemed to be true matches. See description of a true match in Section V.
- I. % of Sample: Percent of records in the random sample with no NDI true match. Column H divided by Column C.
- J. Number of Records with True Match: The number of records for which the NDI returned possible matches and the highest-ranking possible match was deemed to be a true match.
- K. % of Sample: Percent of records in the random sample with an NDI true match. Column J divided by Column C.

Table 11: Agreement Rates with NDI Date of Death for Cases with Activity after Death

| Number of Days After Death | Source of Date of Death | | | |
|--|-------------------------|----------|----------|--------|
| | SSA | Medicare | BIRLS-DF | MSID |
| Number of Deaths in Source | 1,805 | 1,573 | 1,643 | 525 |
| % With Death Date Matching the NDI Death Date | | | | |
| 1-31 | 76.6% | 89.0% | 92.4% | 95.2% |
| 32-61 | 95.7% | 96.3% | 96.4% | 95.1% |
| 62-180 | 94.5% | 94.2% | 96.6% | 96.8% |
| 180-365 | 94.0% | 97.8% | 96.4% | 97.7% |
| 366+ | 95.7% | 95.8% | 96.3% | 97.3% |
| Average for the Sample | 92.9% | 94.9% | 96.0% | 96.6% |
| % With Year and Month of Death Matching the NDI Year and Month of Death | | | | |
| 1-31 | 97.4% | 97.7% | 99.4% | 100.0% |
| 32-61 | 99.4% | 100.0% | 98.7% | 100.0% |
| 62-180 | 99.5% | 99.6% | 99.3% | 99.5% |
| 180-365 | 99.4% | 100.0% | 99.3% | 100.0% |
| 366+ | 100.0% | 100.0% | 99.6% | 100.0% |
| Average for the Sample | 99.3% | 99.6% | 99.3% | 99.8% |

Row Descriptions:

Number of Deaths in Source: Number of death dates found in the source for the veterans with Medicare or VHA activity after their date of death.

1-31, 32-61, etc: Number of days between the most recent date of death in any source and the most recent activity in Medicare or the VHA.

Average for the Sample: This average is a raw average; it is not weighted by probability of selection.

Column Descriptions:

SSA: Source of death date is the SSA Death Master File.

Medicare: Source of death date is the Medicare Vital Status File.

BIRLS-DF: Source of death date is the BIRLS Death File.

MSID: Source of death date is the MSID.

Table 12: Overall Sample Categories and NDI Deaths Identified

| A | B | C | D | E |
|-----------------|--------------------|-------------|------------|---------------|
| Category | Number Of Veterans | % of Sample | NDI Deaths | % of Category |
| Presumed Living | 2,203 | 73.5% | 3 | 0.1% |
| Unknown Status | 503 | 16.8% | 2 | 0.4% |
| Deceased | 292 | 9.7% | 287 | 98.3% |
| Total Sample | 2,998 | 100.0% | 292 | 9.7% |

Row Descriptions:

Presumed Living: The veterans in the Overall Sample who were presumed living as of December 31, 2002. See criteria for this category in Section V.

Unknown Status: Veterans in the Overall Sample who were not presumed living or identified as deceased.

Deceased: Veterans in the Overall Sample who had a date of death in any of the sources (Medicare, SSA, BIRLS-DF or MSID).

Total Sample: The number of veterans in the Overall Sample whose search record was not rejected by the NDI.

Column Descriptions:

A. See Row Descriptions.

B. Number of Veterans: The number of veterans in the categories identified in Column A.

C. % of Sample: Percent of Total Sample.

D. NDI Deaths: The number of deaths identified in the NDI search.

E. % of Category: Percent of the veterans with a date of death in the NDI for the category. Column D divided by Column C.

Table 13: Comparison of Deaths in the Four Sources (BIRLS-DF, SSA, Medicare, MSID) Available Within the VA and the NDI

| VA Sources | NDI Search Results | | | |
|-----------------------|--------------------|---------------------------|---------------------------|-------|
| | Total | Deceased | Living | |
| Deceased | 287 | 5 | 292 | |
| Living | 5 | 2,701 | 2,706 | |
| Total | 292 | 2,706 | 2,998 | |
| Measures of Agreement | | | | |
| Sensitivity | Specificity | Positive Predictive Value | Negative Predictive Value | Kappa |
| 98.3 | 99.8 | 98.3 | 99.8 | 98.1 |

Table 14: NDI Search Results for the Overall Sample

| | Source of Date of Death | | | | |
|--|-------------------------|-------|----------|--------|---------|
| | Medicare | SSA | BIRLS-DF | MSID | Routine |
| Number of Deaths in Source | 243 | 269 | 226 | 35 | 287 |
| % of the 292 NDI Deaths found in a source | 83.2% | 92.1% | 77.4% | 12.0% | 98.3% |
| % With Death Date Agreeing with the NDI Death Date | 96.3% | 95.9% | 97.3% | 100.0% | 98.3% |
| % With Death Date within 2 Days of the NDI Death Date | 97.9% | 98.1% | 98.2% | 100.0% | 99.3% |
| % With Year and Month of Death Agreeing with the NDI Year and Month of Death | 99.6% | 99.6% | 98.7% | 100.0% | 100.0% |

Row Descriptions:

Number of Deaths in Source: Number of death dates found in the source for the veterans in the Overall Sample.

% of the 292 NDI Deaths found in a source: Percent of the 292 deaths identified by the NDI search that are also found in the source.

% With Death Date Agreeing with the NDI Death Date: Percent of the source death dates that agree with the NDI death date.

% With Death Date within 2 Days of the NDI Death Date: Percent of the source death dates that are within 2 days of the NDI death date.

% With Year and Month of Death Agreeing with the NDI Year and Month of Death: Percent of the source death dates that agree with NDI death date on month and year.

Column Descriptions:

Medicare: Source of death date is the Medicare Vital Status File.

SSA: Source of death date is the SSA Death Master File.

BIRLS-DF: Source of death date is the BIRLS Death File.

MSID: Source of death date is the MSID.

Routine: Date is selected from one of the four sources (SSA, Medicare, BIRLS-DF, MSID) based on an automated routine.

Appendix A: Files Used to Create the Study Cohort

| Description | Source File Names | Number of Unique SSNs |
|---|---|------------------------------|
| Users of the VHA Fiscal Years 1997-2002 | RMTPRD.MED.SAS.UNIQSSN.BYSTA3N.FY**.DATA ** = Fiscal years '97' – '03' | 7,523,103 |
| Veterans Enrolled in the VHA 1999-2002 | RMTPRD.R10ARE1.SAS.ENRL1226.OPP9.PSSG RMTPRD.R10ARE1.SAS.ENRL1226.OPP0.PSSG RMTPRD.R10ARE1.SAS.NED.MAIN.DEC01.PSSG RMTPRD.R10ARE1.SAS.NED.MAIN.DEC02.PSSG | 8,849,926 |
| Compensation and Pension Recipients not using the VHA | RMTPRD.MED.SAS.KLFMENU.CPNON.DATA Note: This is a point in time file. This file contained veterans receiving benefits as of July 2003. | 712,811 |
| Compensation and Pension Historical | Special Extract from the VBA of veterans receiving compensation and pension benefits prior to December 2002 if benefits were not terminated before December 2000. | 1,596,549 |
| Pharmacy Only Users 1999-2002 | Extract Provided by ARC | 445,344 |
| PBM 1999 Users | Extract Provided by PBM | 2,749,419 |
| 1999 VIReC Medicare Data Merge Finder File | RMTPRD.R10ARE1.SAS.ENRL1226.OPP9.PSSG RMTPRD.MED.SAS.UNIQSSN.BYSTA3N.FY**.DATA ** = Fiscal years '97' – '00' RMTPRD.MED.SAS.KLFMENU.CPNON.DATA as of April 2000. Extract Provided by ARC of Pharmacy Only Users 1997-1999 | 6,411,210 |
| Total Unique SSNs after merging all the above files | | 9,422,337 |

Appendix B: Files Used to Obtain Demographics for an SSN

| Description | File Name |
|---|--|
| Users of the VHA Fiscal Years 1997-2003 | RMTPRD.MED.SAS.UNIQSSN.BYSTA3N.FY**.DATA ** = Fiscal years '97' – '03' |
| Veterans Enrolled in the VHA 1999-2003 | RMTPRD.R10ARE1.SAS.ENRL1226.OPP9.PSSG RMTPRD.R10ARE1.SAS.ENRL1226.OPP0.PSSG RMTPRD.R10ARE1.SAS.NED.MAIN.DEC01.PSSG RMTPRD.R10ARE1.SAS.NED.MAIN.DEC02.PSSG RMTPRD.R10ARE1.SAS.NED.MAIN.DEC03.PSSG |
| Compensation and Pension Recipients Historical | Special Extracts from the VBA of veterans receiving compensation and pension benefits prior to December 2003 if benefits were not terminated before December 2000. |
| Pharmacy Only Users 1999-2003 | Extract Provided by ARC |
| VBA demographics for Medicare Data Merge Project 2003 Finder File | Special Extract from the VBA of demographics obtained from Compensation and Pension files or BIRLS. |
| Compensation and Pension Recipients not using the VHA | RMTPRD.MED.SAS.KLFMENU.CPNON.DATA Note: This is a point in time file. This file contained veterans receiving benefits as of July 2003. |

Reference List

- (1) Boyle CA, Decoufle P. National sources of vital status information: extent of coverage and possible selectivity in reporting. *Am J Epidemiol* 1990 January;131(1):160-8.
- (2) Cowper DC, Kubal JD, Maynard C, Hynes DM. A primer and comparative review of major US mortality databases. *Ann Epidemiol* 2002 October;12(7):462-8.
- (3) Dominitz JA, Maynard C, Boyko EJ. Assessment of vital status in Department of Veterans Affairs national databases. comparison with state death certificates. *Ann Epidemiol* 2001 July;11(5):286-91.
- (4) Fisher SG, Weber L, Goldberg J, Davis F. Mortality ascertainment in the veteran population: alternatives to the National Death Index. *Am J Epidemiol* 1995 February 1;141(3):242-50.
- (5) Hill ME, Rosenwaike I. The Social Security Administration's Death Master File: the completeness of death reporting at older ages. *Soc Secur Bull* 2001;64(1):45-51.
- (6) Lash TL, Silliman RA. A comparison of the National Death Index and Social Security Administration databases to ascertain vital status. *Epidemiology* 2001 March;12(2):259-61.
- (7) Wentworth DN, Neaton JD, Rasmussen WL. An evaluation of the Social Security Administration master beneficiary record file and the National Death Index in the ascertainment of vital status. *Am J Public Health* 1983 November;73(11):1270-4.
- (8) Maynard C, Chapko MK. Data resources in the Department of Veterans Affairs. *Diabetes Care* 2004 May;27 Suppl 2:B22-B26.
- (9) *VIREC Research User Guide: FY2002 VHA Medical SAS® Inpatient Datasets*. Edward J. Hines, Jr. VA Hospital, Hines, IL: Veterans Affairs Information Resource Center; 2003.
- (10) Bilgrad R. *National Death Index User's Manual*. National Center for Health Statistics; 2000.
- (11) Social Security Number Verification. *Social Security Administration* 2005 January 21; Available at: <http://www.ssa.gov/employer/ssnv.htm>. Accessed March 13, 2006.
- (12) VBA BIRLS Death File Updated. Data Issues Brief October 2003. Veterans Administration Information Resource Center (VIREC).
- (13) *VIREC Research User Guide: FY2002 VHA Medical SAS® Inpatient Datasets*. Edward J. Hines, Jr. VA Hospital, Hines, IL: Veterans Affairs Information Resource Center; 2003.

- (14) Ways and Means Social Security Subcommittee Committee on Financial Services Subcommittee on Oversight and Investigations November 8, 2001. *Social Security Administration* 2005 January 25; Available at: http://www.ssa.gov/legislation/testimony_110801.html. Accessed March 13, 2006.
- (15) Use of Social Security Administration Data for Research Purposes. Social Security Administration; 2004. Report No.: 2.
- (16) Doody MM, Chimes K. The Social Security Administration "presumed living" search. *Am J Public Health* 2000 December;90(12):1948-9.
- (17) Fleming C, Fisher ES, Chang CH, Bubolz TA, Malenka DJ. Studying outcomes and hospital utilization in the elderly. The advantages of a merged data base for Medicare and Veterans Affairs hospitals. *Med Care* 1992 May;30(5):377-91.