



Practice of Epidemiology

Utility of the National Death Index in Ascertaining Mortality in Acquired Immunodeficiency Syndrome Surveillance

Mary Jo Trepka*, Lorene M. Maddox, Spencer Lieb, and Theophile Niyonsenga

* Correspondence to Dr. Mary Jo Trepka, Department of Epidemiology and Biostatistics, Robert Stempel College of Public Health and Social Work, Florida International University, University Park, HLS II, Room 595, 11200 Southwest 8th Street, Miami, FL 33199 (e-mail: trepkam@fiu.edu).

Initially submitted October 25, 2010; accepted for publication January 24, 2011.

To assess the utility of the National Death Index (NDI) in improving the ascertainment of deaths among people diagnosed with acquired immunodeficiency syndrome (AIDS), the authors determined the number and characteristics of additional deaths identified through NDI linkage not ascertained by using standard electronic linkage with Florida Vital Records and the Social Security Administration's Death Master File. Records of people diagnosed with acquired immunodeficiency syndrome between 1993 and 2007 in Florida were linked to the NDI. The demographic characteristics and reported human immunodeficiency virus (HIV) transmission modes of people whose deaths were identified by using the NDI were compared with those whose deaths were ascertained by standard linkage methods. Of the 15,094 submitted records, 719 had confirmed matches, comprising 2.1% of known deaths ($n = 34,504$) within the cohort. Hispanics, males, people 40 years of age or older, and injection drug users were overrepresented among deaths ascertained only by the NDI. In-state deaths comprised 59.0% of newly identified deaths, and human immunodeficiency virus was less likely to be a cause of death among newly identified compared with previously identified deaths. The newly identified deaths were not previously ascertained principally because of slight differences in personal identifying information and could have been identified through improved linkages with Florida Vital Records.

acquired immunodeficiency syndrome; mortality; population surveillance; registries; vital statistics

Abbreviations: AIDS, acquired immunodeficiency syndrome; CDC, Centers for Disease Control and Prevention; CI, confidence interval; HARS, HIV/AIDS Reporting System; HIV, human immunodeficiency virus; NDI, National Death Index; SSDMF, Social Security Administration's Death Master File.

Complete ascertainment of deaths among people in human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS) surveillance systems is necessary for the accurate calculation of prevalence rates and mortality patterns among people living with HIV/AIDS. These data are used to monitor the epidemic and in formulas to distribute federal funding for the care and support of people living with HIV/AIDS (1). Although in-state deaths can be ascertained by each state linking its vital records data with data from HIV/AIDS surveillance systems, also known as the HIV/AIDS Reporting System (HARS), out-of-state deaths must generally be ascertained by electronic linkage with the Social Security Administration's Death Master File (SSDMF) or the National Death Index (NDI). The Centers

for Disease Control and Prevention (CDC) facilitates the linkage of state HARS data with the SSDMF by providing states with a copy of the SSDMF and Link Plus, a linkage program developed by the CDC (2).

The SSDMF is a database of deaths reported to the Social Security Administration among people with a Social Security number (3). Linkage with the SSDMF results in the ascertainment of deaths for people who die outside of a given state or in a foreign country provided the death was reported to the Social Security Administration. Nevertheless, data indicate that 10%–25% of deaths occurring since 1990 are not captured by using the SSDMF (3, 4). Furthermore, the sensitivity of the SSDMF may be lower for deaths occurring among people in HARS because the SSDMF is

less sensitive among younger age groups (e.g., 74% among those aged 25–45 years vs. >95% among those aged 65 years or older) (3) and the unmarried (5) and may be lower among hard-to-reach populations such as sex workers (6).

The NDI is an electronic registry of death record data obtained from the vital statistics offices of all states (7). Thus, the NDI should be a complete source of deaths for people in HARS who died in the United States. However, use of the NDI is costly because of the linkage charge and the considerable staff time to review each match's validity. Studies have found the NDI's sensitivity to range from 93% to 100% (8–14). There has been 1 published study that evaluated the benefit of the SSDMF and the NDI for HIV/AIDS surveillance (15). In this study, records from the New York City HIV registry were matched with the NDI and the SSDMF, and each source was found to be the sole source for some deaths.

To ascertain the vital status of people in the Florida HARS, the Florida Department of Health Bureau of HIV/AIDS has been linking Florida HARS data with Florida Vital Records and the SSDMF since 2006. However, the Florida HARS was not linked with the NDI since 1993 because of cost. In the fall of 2009, funding became available for linking AIDS cases in HARS with the NDI. To evaluate the utility of the NDI in improving the ascertainment of deaths, the current study was conducted to determine 1) the number of additional deaths identified through NDI linkage that were not ascertained by using the standard electronic linkage with Florida Vital Records and the SSDMF and 2) if there are subgroups of people whose deaths are less likely to be ascertained by using the standard linkage.

MATERIALS AND METHODS

Ascertaining vital status of Florida HIV/AIDS Reporting System cases

The vital status of cases in the HARS (evolved into eHARS in 2009 but will be referred to as "HARS" hereafter for simplicity) is ascertained through linkage with the Florida Office of Vital Statistics records and the SSDMF. Death certificates coded with HIV-related *International Classification of Diseases*, Ninth Revision (prior to 1999) and Tenth Revision (1999 and later), death codes are automatically submitted electronically to HARS staff as soon as the deaths are reported to the Florida Office of Vital Statistics, and HARS data are linked quarterly to the state vital records electronic database. Currently, these linkages are based only on an exact match of the first name, last name, Social Security number, and date of birth.

Since 2006, the CDC has annually provided the Bureau of HIV/AIDS with an updated SSDMF database for linkage with HARS data. The primary variables used for linkage are last name, first name, Social Security number, and date of birth. These deaths are linked by using Registry Plus Link Plus 2.012 (CDC, Atlanta, Georgia). The Link Plus software gives a score based on the probability of a match. The highest score is given for exact matches on spelling of first and last names, Social Security number, and date of birth. The score drops with any slight variation including phonetic

matches and drops further if fewer than 4 variables match. Zip code, county, and any alias for name, date of birth, or Social Security number are used to validate matches. Staff members review the lowest scoring matches first and discard validated nonmatches. Probable matches are validated by reviewing an electronic database of all scanned documents associated with cases, the Florida Vital Records Registry, and the Judicial Inquiry System to see if there were any documented court issues that might identify a name change or alternative date of birth or Social Security number.

National Death Index linkage

Funding for NDI linkage was sufficient only for AIDS cases that were presumed living; thus, HIV (not AIDS) cases were not submitted for matching. After approval was obtained from the National Center for Health Statistics for linkage in October 2009, the data set was prepared. At the time of the linkage, the NDI had death certificate data through December 2007. To decrease the number of records to submit for linkage, records with a laboratory report date or date of Medicaid, Florida AIDS Drug Assistance Program, or health department HIV-related care after December 2007 were deleted. The variables submitted to the NDI for matching included last name, first name, middle initial, Social Security number, date of birth, sex, race, current state of residence, and state of birth. Two additional data sets were sent. One included cases in the first data set with an alias Social Security number, and the other included cases with an alias date of birth.

For confidentiality reasons, the NDI does not return names or numbers relating to a "possible" match; instead, it codes an exact match with an "X" for that variable in the output from the NDI linkage. Deaths with partial or similar matches on names were indicated by an "I" (for agreement on first initial) or an "N" (for presumed phonetic agreement). The Social Security number was matched digit for digit. (If all digits match, there would be 9 X's; if any digit didn't match, the digit was coded by a dash.) The date of birth was matched on month, day, and year. If the year of birth was not an exact match, the output showed the number of years (+ or -) from the NDI record. A more detailed description of the linkage process and generated output can be found on the NDI Web site (7).

The first step of the match validation was conducted by 1 primary reviewer and involved triaging the possible matches to determine the probability of a true match. If there was only a partial match on name and Social Security number and the year of death was before dates for when the case was known to be alive on the basis of clinical or laboratory records, the match was discarded from further review. In addition, matches with low levels of agreement (e.g., only first name and sex, only first name and race, only last name and sex, only last name and race, Social Security number matched on fewer than 5 digits and year of birth off by more than 5 years) between the NDI and HARS were bypassed from further review. The reviewer conducted detailed reviews for all other cases by first reviewing an electronic database of all scanned documents associated with cases. For out-of-state deaths, the reviewer searched the database

for any history of that case's association with that state, such as documented phone calls. If the person was known to have an incarceration history, the reviewer also searched the Florida Department of Corrections database to locate the Social Security number or other identifying information that might not be in HARS. The reviewer searched the SSDMF to determine if the death was identified in that database. Previous linkages with the SSDMF may have missed the death if it didn't match on full name, date of birth, and Social Security number, especially for the people for whom we had no Social Security number or date of birth on file in our database to match. Because the SSDMF database provides the actual name, date of birth, and Social Security number, a match in the SSDMF database confirmed the coded match found via the NDI. To determine how sensitive the first reviewer was at identifying "probable" matches, a second reviewer performed the same procedure for a 5% sample of records.

The matches deemed "probable" during the initial detailed review were divided between 2 secondary reviewers. The secondary reviewers looked for the cases in the Judicial Inquiry System. If the death was in Florida, the reviewers searched Florida Vital Records to identify the actual death certificate.

Statistical analysis

The distributions of demographic and mode-of-transmission characteristics were compared between deaths ascertained with the usual state vital records and SSDMF linkages and those ascertained by the NDI by calculating the crude and adjusted odds ratios using simple and multiple logistic regression. Variables in the model included year of death, sex, age group at time of death, race/ethnicity, country of birth, and mode of HIV transmission. For country-of-birth classification, people were classified as being foreign born if they were not born in a state or dependency of the United States. For the HIV/AIDS transmission mode, cases with a combined reported transmission mode of men who have sex with men and injection drug use were grouped with cases with the injection-drug-use transmission mode. Deaths with an *International Classification of Diseases*, Ninth Revision, code of 042-044 or an *International Classification of Diseases*, Tenth Revision, code of B20-B24 were classified as having an underlying cause of death of HIV. The AIDS prevalence rate was calculated with and without the NDI results. Statistical significance was indicated if $P < 0.05$. Analyses were conducted by using Statistical Analysis System, version 9.1, software (SAS Institute, Inc., Cary, North Carolina). The Florida Department of Health Institutional Review Board approved this study, and the Florida International University Institutional Review Board deemed the current study exempt from institutional review board approval.

RESULTS

There were a total of 77,186 people reported with AIDS between 1993 and 2007. Of these, 33,785 had a death record in the HARS database as of December 2009, leaving 43,401

people presumed living (Figure 1). For 28,307 (65.2%) of these presumed living people, there was a date for a laboratory result or HIV-related care after December 2007. Therefore, these people were confirmed alive as of December 2007, and their records were deleted from the NDI linkage dataset, leaving 15,094 cases. Of these, 7,694 (51.0%) records had "possible" matches, meaning they agreed on 1 or more variables in HARS and the NDI. Cases with common names had more possible matches. All submitted variables agreed on 126 (1.6%) records. After the initial triage and detailed review, 773 of the possible matches were deemed "probable" matches, including several questionable matches that were not in the SSDMF database, matches for people who died out of state, or matches for people who had a common name but not exact (but close) match on date of birth or Social Security number. Most of the deaths that were identified through the NDI and found only in the SSDMF after NDI linkage were cases with a missing or incorrect Social Security number in HARS and/or with some other variable such as first name, last name, or part of date of birth that did not match. A second review of a 5% sample of possible matches failed to identify any additional probable matches.

Of the 773 probable matches, 719 deaths were confirmed as matches during the subsequent detailed review. These comprised 2.1% (95% confidence interval (CI): 1.9, 2.2) of all known deaths ($n = 34,504$) within the cohort. Recalculating the AIDS prevalence rate in 2007 by excluding from the numerator cases ascertained only by NDI linkage with a death prior to 2007 decreased the AIDS prevalence rate in Florida by 1.5% from 261.28 to 257.44 cases per 100,000. Relative to deaths ascertained by Florida Vital Records or an SSDMF match, deaths occurring after 1995 and among people who were male, ≥ 40 years of age, and Hispanic were overrepresented among deaths ascertained only by the NDI (Table 1). Deaths among people with a reported HIV transmission mode of injection drug use were overrepresented among deaths ascertained only by NDI linkage relative to men who have sex with men. Almost all the deaths prior to the era of highly active antiretroviral therapy, which began in 1996–1997, were ascertained by Florida Vital Records or SSDMF (99.3%). Subsequently, during a period when overall AIDS deaths in Florida and the United States were declining (data not shown), the percentage ascertained only from the NDI increased. Of the 34,504 deaths, 32,151 (93.2%) were within the state of Florida, 2,232 (6.5%) were in another state or US dependency, and 121 (0.4%) were in another country or had a missing location of death. The deaths outside of Florida were distributed among 48 states and Puerto Rico with the largest number occurring in Georgia ($n = 307$) and New York ($n = 307$). Of the 719 deaths ascertained only by the NDI, 424 (59.0%) were within the state of Florida, and 295 (41.0%) were outside of Florida. For deaths within the state of Florida, 1.3% (95% CI: 1.2, 1.4) were ascertained only by the NDI. For deaths occurring outside the state of Florida, 13.2% (95% CI: 11.8, 14.6) were ascertained only by the NDI.

Of deaths that occurred within the state of Florida, deaths occurring after 1995 and deaths among males, non-Hispanic

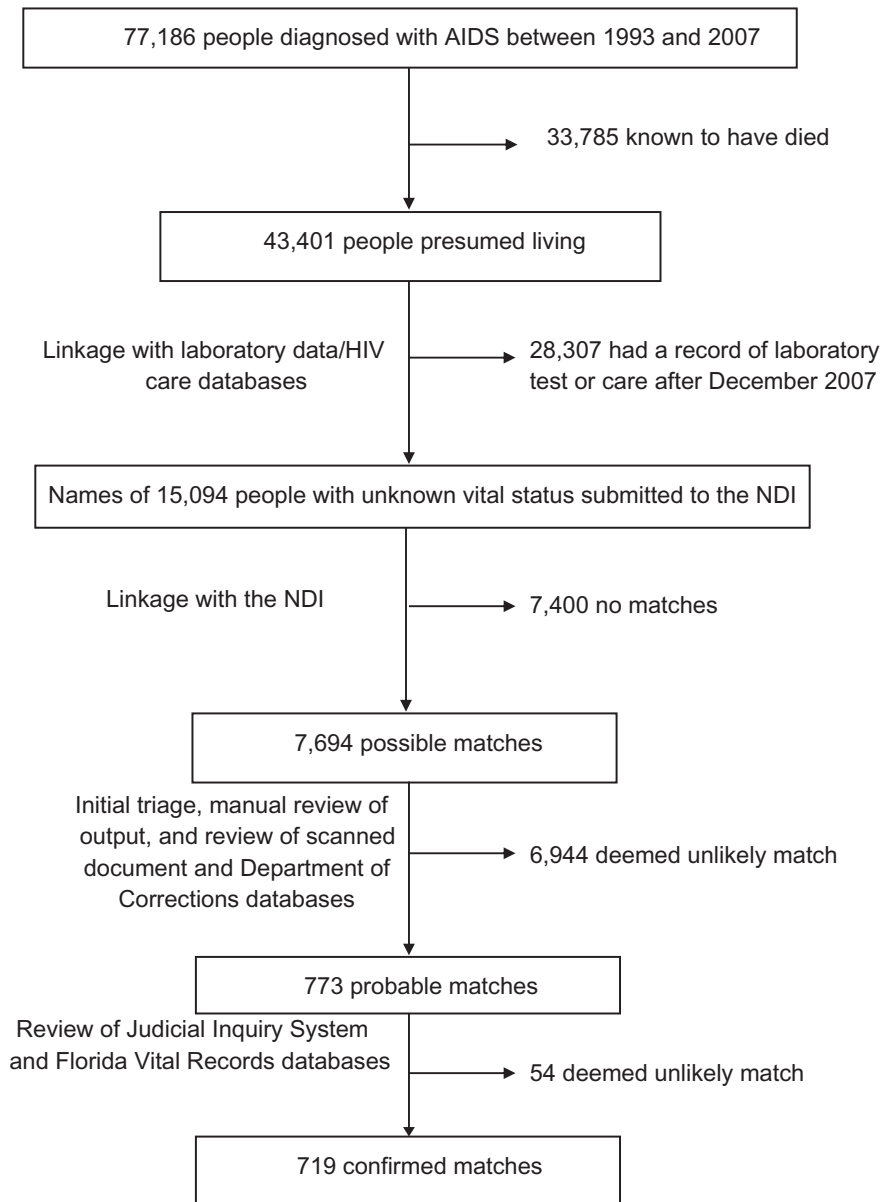


Figure 1. Flowchart of linkage of HIV/AIDS Reporting System records of people diagnosed with AIDS in Florida from 1993 to 2007 with the National Death Index. AIDS, acquired immunodeficiency syndrome; HIV, human immunodeficiency virus; NDI, National Death Index.

blacks, Hispanics, and foreign born were overrepresented among deaths ascertained only by the NDI (Table 2). Deaths among people with a reported HIV transmission mode of injection drug use were overrepresented relative to men who have sex with men. In addition, older age groups were more likely to be ascertained only by the NDI than younger age groups, and the association strengthened the older the age group. For 27,201 (84.6%) of Florida deaths, there was an underlying cause of death. Of these Florida deaths, the cause of death was HIV among 14.1% of deaths ascertained only by the NDI compared with 62.1% of deaths identified by standard linkage methods (chi-square, $P < 0.0001$).

Among the deaths in other US states and Puerto Rico, those occurring after 1995 and deaths among Hispanics and non-Hispanic blacks were overrepresented among deaths ascertained only by the NDI (Table 3). Cause of death could not be analyzed for out-of-state deaths because of missing cause-of-death data for 81.5% of deaths.

For in-state (Table 2) and out-of-state (Table 3) deaths but not for all deaths combined (Table 1), black race was associated with being overrepresented among deaths ascertained only by the NDI. For in-state (Table 2) but not out-of-state (Table 3) deaths, males, injection drug users relative to men who have sex with men, people older

Table 1. Distribution of Deaths Among People Diagnosed With Acquired Immunodeficiency Syndrome in Florida from 1993 to 2007 by Source of How Death Information Was Ascertained

Characteristic	No. of Deaths	% Ascertained by Florida Vital Records or the SSDMF ^a	% Ascertained Only From the NDI ^a	Crude OR	95% CI	Adjusted OR	95% CI
Total	34,504	97.9	2.1				
Year of death							
1993–1995	7,536	99.3	0.7	1.00	Referent	1.00	Referent
1996–1998	7,354	98.5	1.6	2.27	1.63, 3.15	2.23	1.61, 3.12
1999–2001	6,255	97.5	2.5	3.75	2.74, 5.14	3.64	2.65, 5.00
2002–2004	6,650	96.1	3.9	5.83	4.32, 7.87	5.67	4.19, 7.68
2005–2007	6,709	98.0	2.0	2.96	2.14, 4.08	2.83	2.04, 3.93
Sex							
Males	25,006	97.9	2.2	1.00	Referent	1.00	Referent
Females	9,498	98.1	1.9	0.88	0.75, 1.05	0.74	0.61, 0.90
Age at time of death, years							
0–29	2,721	98.9	1.1	1.00	Referent	1.00	Referent
30–39	10,430	98.2	1.8	1.58	1.08, 2.31	1.42	0.97, 2.09
40–49	11,859	97.8	2.2	1.98	1.36, 2.88	1.50	1.02, 2.19
50–59	6,021	97.5	2.5	2.25	1.52, 3.32	1.61	1.08, 2.40
≥60	3,473	97.6	2.5	2.18	1.44, 3.29	1.53	1.01, 2.34
Race/ethnicity							
Non-Hispanic white	9,618	98.3	1.7	1.00	Referent	1.00	Referent
Non-Hispanic black	18,987	97.9	2.2	1.26	1.05, 1.51	1.06	0.87, 1.30
Hispanic	5,198	97.4	2.7	1.56	1.24, 1.96	1.42	1.10, 1.83
Other/unknown	701	98.9	1.1	0.66	0.32, 1.35	0.54	0.26, 1.10
Country of birth							
Foreign born/unknown	6,534	97.6	2.4	1.00	Referent	1.00	Referent
US born ^b	27,970	98.0	2.0	0.85	0.71, 1.02	0.94	0.77, 1.16
Mode of HIV transmission							
Injection drug use ^c	7,343	97.5	2.5	1.00	Referent	1.00	Referent
Men who have sex with men	11,263	98.4	1.6	0.64	0.52, 0.79	0.62	0.50, 0.77
Heterosexual contact	9,747	97.8	2.2	0.89	0.73, 1.09	0.90	0.73, 1.12
Other/unknown	6,151	97.8	2.2	0.87	0.69, 1.09	0.89	0.70, 1.12

Abbreviations: CI, confidence interval; HIV, human immunodeficiency virus; NDI, National Death Index; OR, odds ratio; SSDMF, Social Security Administration's Death Master File.

^a Percentage totals may exceed 100% because of rounding.

^b "US born" includes people born in any US state or dependency.

^c "Injection drug use" includes men who reported having sex with men if they also reported injection drug use as mode of transmission.

than 50 years, and the foreign born were overrepresented among deaths ascertained only by the NDI.

The cost of the linkage with the NDI was \$21,749. The cost for staff time including salary and fringe benefits was estimated at \$7,200–\$9,600 (representing 300–400 hours) for a total cost of \$28,949–\$31,349 or \$40.26–\$43.60 per additional case ascertained. The linkage and validation process took 5 months.

DISCUSSION

The principal finding in this study is that the ongoing linkages with Florida Vital Records and the SSDMF

identified the vast majority of deaths, with subsequent NDI linkage responsible for only 2.1% of all deaths ascertained. This is similar to the finding from the HIV surveillance study in New York City, in which only 3.2% of all deaths were found by NDI linkage alone (15). However, there were several important lessons learned from the deaths identified only by the NDI. First, contrary to expectations, most deaths identified only by the NDI occurred in Florida, despite regular linkages to Florida Vital Records data. Second, the demographic and mode-of-transmission characteristics of people whose deaths were identified by the NDI differed significantly from those whose deaths were found by other methods.

Table 2. Distribution of Deaths Occurring in Florida Among People Diagnosed With Acquired Immunodeficiency Syndrome in Florida From 1993 to 2007 by Source of How Death Information Was Ascertained

Characteristic	No. of Deaths	% Ascertained by Florida Vital Records or the SSDMF ^a	% Ascertained Only From the NDI ^a	Crude OR	95% CI	Adjusted OR	95% CI
Total	32,151	98.7	1.3				
Year of death							
1993–1995	7,109	99.7	0.3	1.00	Referent	1.00	Referent
1996–1998	6,888	99.1	0.9	2.71	1.67, 4.38	2.58	1.59, 4.17
1999–2001	5,791	98.2	1.8	5.52	3.51, 8.70	4.97	3.15, 7.84
2002–2004	6,175	97.4	2.6	8.09	5.22, 12.55	7.17	4.60, 11.17
2005–2007	6,188	98.7	1.3	4.09	2.57, 6.50	3.50	2.18, 5.60
Sex							
Males	23,073	98.6	1.4	1.00	Referent	1.00	Referent
Females	9,078	98.8	1.2	0.89	0.72, 1.11	0.70	0.55, 0.89
Age at time of death, years							
0–29	2,577	99.3	0.7	1.00	Referent	1.00	Referent
30–39	9,600	99.2	0.9	1.30	0.77, 2.19	1.21	0.71, 2.05
40–49	10,972	98.6	1.4	2.07	1.25, 3.43	1.59	0.95, 2.65
50–59	5,664	98.2	1.8	2.76	1.65, 4.62	1.97	1.17, 3.33
≥60	3,338	97.8	2.2	3.41	2.01, 5.80	2.23	1.30, 3.83
Race/ethnicity							
Non-Hispanic white	8,471	99.3	0.7	1.00	Referent	1.00	Referent
Non-Hispanic black	18,212	98.4	1.6	2.18	1.65, 2.88	1.59	1.18, 2.15
Hispanic	4,810	98.5	1.5	2.15	1.53, 3.03	1.55	1.07, 2.25
Other/unknown	658	99.2	0.8	1.06	0.42, 2.64	0.74	0.30, 1.87
Country of birth							
Foreign born/unknown	6,276	98.1	1.9	1.00	Referent	1.00	Referent
US born ^b	25,875	98.9	1.2	0.63	0.50, 0.77	0.69	0.54, 0.88
Mode of HIV transmission							
Injection drug use ^c	6,624	98.6	1.4	1.00	Referent	1.00	Referent
Men who have sex with men	10,257	99.2	0.8	0.56	0.41, 0.75	0.56	0.41, 0.77
Heterosexual contact	9,348	98.4	1.6	1.15	0.88, 1.49	1.00	0.76, 1.32
Other/unknown	5,922	98.3	1.7	1.21	0.91, 1.60	1.04	0.77, 1.40

Abbreviations: CI, confidence interval; HIV, human immunodeficiency virus; NDI, National Death Index; OR, odds ratio; SSDMF, Social Security Administration's Death Master File.

^a Percentage totals may exceed 100% because of rounding.

^b "US born" includes people born in any US state or dependency.

^c "Injection drug use" includes men who reported having sex with men if they also reported injection drug use as mode of transmission.

The questions arise why most deaths within Florida and identified only by the NDI were not identified through regular linkages with Florida Vital Records, and why Hispanics, non-Hispanic blacks, foreign born, men, and injection drug users were overrepresented among the previously unknown Florida deaths. The explanation is probably due to linkages being based on exact matches, which require that both the death certificate and HARS have the exact same Social Security number, date of birth, and spelling of the first and last name for a given person. Many of these deaths were among people with missing or inaccurate personal identifying information such as Social Security number or date of birth. If the linkage methods with vital records could be improved to provide for possible matches similar to

those provided by NDI linkages, many of the missing deaths would be identified. This would be less costly than using NDI because there would be no linkage fees. The high proportion of deaths identified only by the NDI with an underlying cause of death that is not HIV relative to the proportion of deaths identified by using standard methods is likely due to the Florida Office of Vital Statistics automatically flagging and reporting deaths with an underlying cause of death of HIV to the Florida Bureau of HIV/AIDS. Therefore, in Florida as well as in other states that have the same procedures for ascertaining deaths in AIDS registries, deaths due to HIV may be overrepresented in cause-of-death analyses unless there is NDI linkage. Females were less likely to be identified by NDI only in the model that adjusted

Table 3. Distribution of Deaths Occurring Outside of Florida but Within the United States or Puerto Rico Among People Diagnosed With Acquired Immunodeficiency Syndrome in Florida from 1993 to 2007 by Source of How Death Information Was Ascertained

Characteristic	No. of Deaths	% Ascertained by the SSDMF ^a	% Ascertained Only From the NDI ^a	Crude OR	95% CI	Adjusted OR	95% CI
Total	2,232	86.8	13.2				
Year of death							
1993–1995	395	92.7	7.3	1.00	Referent	1.00	Referent
1996–1998	446	87.9	12.1	1.74	1.08, 2.79	1.63	1.01, 2.63
1999–2001	442	87.1	12.9	1.87	1.17, 2.99	1.65	1.02, 2.65
2002–2004	452	77.7	22.4	3.63	2.34, 5.63	3.20	2.04, 5.02
2005–2007	497	89.1	10.9	1.54	0.96, 2.47	1.39	0.85, 2.26
Sex							
Males	1,840	87.8	12.2	1.00	Referent	1.00	Referent
Females	392	81.9	18.1	1.60	1.19, 2.14	1.26	0.88, 1.79
Age at time of death, years							
0–29	131	89.3	10.7	1.00	Referent	1.00	Referent
30–39	796	86.9	13.1	1.26	0.70, 2.27	1.37	0.74, 2.51
40–49	843	86.2	13.8	1.33	0.74, 2.40	1.34	0.72, 2.47
50–59	336	85.1	14.9	1.46	0.78, 2.75	1.42	0.74, 2.75
≥60	126	91.3	8.7	0.80	0.35, 1.83	0.69	0.29, 1.63
Race/ethnicity							
Non-Hispanic white	1,132	90.8	9.2	1.00	Referent	1.00	Referent
Non-Hispanic black	732	83.1	16.9	2.02	1.53, 2.67	1.72	1.27, 2.33
Hispanic	325	80.3	19.7	2.42	1.73, 3.40	2.20	1.52, 3.18
Other/unknown	43	93.0	7.0	0.74	0.23, 2.44	0.69	0.21, 2.32
Country of birth							
Foreign born/unknown	201	82.1	17.9	1.00	Referent	1.00	Referent
US born ^b	2,031	87.3	12.8	0.67	0.46, 0.98	0.87	0.57, 1.32
Mode of HIV transmission							
Injection drug use ^c	683	86.7	13.3	1.00	Referent	1.00	Referent
Men who have sex with men	982	89.6	10.4	0.75	0.56, 1.02	1.00	0.72, 1.39
Heterosexual contact	368	81.5	18.5	1.48	1.05, 2.08	1.31	0.90, 1.92
Other/unknown	199	82.9	17.1	1.34	0.87, 2.06	1.41	0.90, 2.23

Abbreviations: CI, confidence interval; HIV, human immunodeficiency virus; NDI, National Death Index; OR, odds ratio; SSDMF, Social Security Administration's Death Master File.

^a Percentage totals may exceed 100% because of rounding.

^b "US born" includes people born in any US state or dependency.

^c "Injection drug use" includes men who reported having sex with men if they also reported injection drug use as mode of transmission.

for mode of HIV transmission but not in the unadjusted analysis. This is likely due to men who have sex with men also being less likely to be identified by NDI only. To determine if the association between gender and identification by NDI only was also confounded by cause of death, we added the cause-of-death variable (HIV vs. all other) to the multivariate model, and the association did not change.

For several factors including age at time of death, year of death, being foreign born, and mode of transmission, there was a relation between being ascertained by NDI only for deaths within Florida but not outside of Florida. Because Florida death certificates with a cause of death of HIV infection are automatically sent to Florida HARS, the cause of death could confound the relation between these factors and

ascertainment of death only by NDI. When cause of death was added to the multivariate model for Florida deaths, the relation between age group and how death was ascertained did indeed weaken but remained significant for those groups aged 50–59 years and 60 years or older. The association with year of death weakened but was still significant for all time periods after 1996. The association with men who have sex with men was virtually unchanged. Therefore, the cause of death may partially explain the association between being identified by NDI only and older age and more recent deaths. However, it does not fully explain the patterns seen, and cause of death does not explain the association with mode of HIV transmission.

The deaths that occurred outside Florida that were identified by the NDI can be interpreted as deaths missed by

SSDMF linkages. That non-Hispanic blacks and Hispanics were overrepresented among these deaths could be due to the specific people not being in the SSDMF at the time of previous linkage or differences in linkage variables between Florida HARS and the SSDMF. Most (90%) deaths that are in the SSDMF are reported to the Social Security Administration by funeral directors, relatives, and friends (16). However, the Social Security Administration does not include a death in the SSDMF if there is a discrepancy between data associated with a particular Social Security number and the death report until the discrepancy is resolved. A 1998 review found that 1.3 million deaths for which the Social Security Administration had terminated benefits based on death information were not included in the SSDMF. Most of these deaths were not included because of discrepancies in name information (62.2%) or date of birth (32.4%) (17). Therefore, it is possible that some of the deaths subsequently identified through the NDI linkage were not in the SSDMF at the time of previous linkages either because no one had reported it to the Social Security Administration or because there were discrepancies in the personal identifying information. The second possibility is that the deaths were in the SSDMF but did not link with Florida HARS because of discrepancies between the Florida HARS and the SSDMF in name, date of birth, or Social Security number.

Linkage with the NDI did not substantially change the overall prevalence rates for people living with AIDS in Florida or in New York City (15). It is likely that this result would vary by state according to linkage methods and the quality of identifying information, such as Social Security number and date of birth. It is also possible that the specific subgroups that are more likely to be missed during standard linkage would vary somewhat by state. In the New York City study (based on calculations of presented data), as in the current study, Hispanics were overrepresented among those deaths identified only by the NDI (15). However, unlike the current study, deaths among younger people were more likely to be identified only by the NDI in New York City. Surveillance staff should draw upon their knowledge of the limitations of their surveillance system data in interpreting vital status data, because mortality may be underestimated in some groups such as injection drug users if they are more likely to have inaccurate or incomplete personal identifying information.

Linkage with the NDI was costly because of the linkage charge and staff time required to review the possible matches. The linkage cost was substantially decreased by reviewing laboratory and clinical records to decrease the number of records that had to be submitted for linkage. The staff time could be decreased somewhat by using an electronic triage system that assigns triage scores based on weighting matching variables (15, 18).

There are some limitations of this study. First, only people with HIV infection who were diagnosed with AIDS were included in this study. It is possible that results could differ for those diagnosed with HIV infection but who have not advanced to AIDS. Second, it is possible that some valid matches were discarded in the initial triage process, although the rereview of the sample of cases indicates that

the number of such matches, if any, would have been small. Third, deaths outside of the United States and its dependencies would not be identified by the NDI, and it is likely that these deaths are also not fully captured by the SSDMF.

Our results provide reassurance that death ascertainment in Florida HARS by using the SSDMF and Florida Vital Records was quite complete. However, death ascertainment was less complete in some populations such as Hispanics, males, and injection drugs users and for those whose underlying cause of death was not HIV. Florida HARS surveillance staff should consider enhancing linkage methods with Florida Vital Records to allow for multiple matches and not only exact matches. There is a 2-year delay between the year of death and inclusion of a death in the NDI. Therefore, even if the NDI becomes available to each state's HARS, each state's vital records data will likely be a more timely source of death data for HARS. Thus, linkage with state vital records should be enhanced whenever possible for HARS and other disease registries for which timeliness of vital status is important. HARS surveillance staff in other states should also consider if inaccuracies in personal identification variables in their registries may be impacting the completeness of death ascertainment in their state, particularly for specific populations.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology and Biostatistics, Robert Stempel College of Public Health and Social Work, Florida International University, Miami, Florida (Mary Jo Trepka, Theophile Niyonsenga); and Bureau of HIV/AIDS, Florida Department of Health, Tallahassee, Florida (Lorene M. Maddox, Spencer Lieb).

The project described was supported by Award R01MD004002 from the National Institute on Minority Health and Health Disparities at the National Institutes of Health.

The authors wish to thank Tracina Bush and Julia Fitz for assistance in reviewing the NDI matches.

The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Minority Health and Health Disparities or the National Institutes of Health.

Conflict of interest: none declared.

REFERENCES

1. About the Ryan White HIV/AIDS Program. Rockville, MD: Health Resources and Services Administration, US Department of Health and Human Services; 2010. (<http://hab.hrsa.gov/aboutus.htm>). (Accessed June 22, 2010).
2. Electronic record linkage to identify deaths among persons with AIDS—District of Columbia, 2000–2005. *MMWR Morb Mortal Wkly Rep.* 2008;57(23):631–634.
3. Hill ME, Rosenwaike I. The Social Security Administration's Death Master File: the completeness of death reporting at older ages. *Soc Secur Bull.* 2001–2002;64(1):45–51.

4. Schnorr TM, Steenland K. Identifying deaths before 1979 using the Social Security Administration Death Master File. *Epidemiology*. 1997;8(3):321–323.
5. 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;73(11):1270–1274.
6. Potterat JJ, Brewer DD, Muth SQ, et al. Mortality in a long-term open cohort of prostitute women. *Am J Epidemiol*. 2004;159(8):778–785.
7. National Center for Health Statistics, Centers for Disease Control and Prevention. National Death Index. Atlanta, GA: Centers for Disease Control and Prevention; 2010. (<http://www.cdc.gov/nchs/ndi.htm>). (Accessed September 9, 2010).
8. Boyle CA, Decouflé P. National sources of vital status information: extent of coverage and possible selectivity in reporting. *Am J Epidemiol*. 1990;131(1):160–168.
9. Calle EE, Terrell DD. Utility of the National Death Index for ascertainment of mortality among Cancer Prevention Study II participants. *Am J Epidemiol*. 1993;137(2):235–241.
10. Davis KB, Fisher L, Gillespie MJ, et al. A test of the National Death Index using the Coronary Artery Surgery Study (CASS). *Control Clin Trials*. 1985;6(3):179–191.
11. Fisher SG, Weber L, Goldberg J, et al. Mortality ascertainment in the veteran population: alternatives to the National Death Index. *Am J Epidemiol*. 1995;141(3):242–250.
12. Kraut A, Chan E, Landrigan PJ. The costs of searching for deaths: National Death Index vs Social Security Administration. *Am J Public Health*. 1992;82(5):760–761.
13. Lash TL, Silliman RA. A comparison of the National Death Index and Social Security Administration databases to ascertain vital status. *Epidemiology*. 2001;12(2):259–261.
14. Stampfer MJ, Willet WC, Speizer FE, et al. Test of the National Death Index. *Am J Epidemiol*. 1984;119(5):837–839.
15. Hanna DB, Pfeiffer MR, Sackoff JE, et al. Comparing the National Death Index and the Social Security Administration's Death Master File to ascertain death in HIV surveillance. *Public Health Rep*. 2009;124(6):850–860.
16. Schaeffer SL. Congressional response report, "The Social Security Administration's efforts to process death reports and improve its Death Master File," A-09-03-23067. Baltimore, MD: Office of the Inspector General, Social Security Administration; 2003: 1–10. (<http://www.ssa.gov/oig/ADOBEPDF/A-09-03-23067.pdf>). (Accessed August 16, 2010).
17. Huse JG Jr. "Improving the usefulness of Social Security Administration's Death Master File, evaluation report," A-09-98-61011. Baltimore, MD: Office of the Inspector General, Social Security Administration; 2000: 1–9. (<http://www.ssa.gov/oig/ADOBEPDF/A-09-98-61011.pdf>). (Accessed August 16, 2010).
18. Fillenbaum GG, Burchett BM, Blazer DG. Identifying a National Death Index match. *Am J Epidemiol*. 2009;170(4):515–518.