COMPARISON OF TWO NATIONAL DEATH DATABASES FOR DEATH ASCERTAINMENT IN PERSONS REPORTED WITH HIV/AIDS IN NEW YORK CITY

David B. Hanna, Melissa R. Pfeiffer, Judith E. Sackoff, Lucia V. Torian

New York City Department of Health and Mental Hygiene, New York, NY, USA

HIV Epidemiology and Field Services Program
Bureau of HIV/AIDS Prevention and Control
New York City Department of Health and Mental Hygiene
346 Broadway, Room 706, CN 44, New York, NY 10013
Phone: +1 212 442 3513; Email: dhanna@health.nyc.gov

BACKGROUND

- Accurate monitoring of the HIV/AIDS epidemic, including the number of persons with HIV/AIDS (PWHA) and deaths to PWHA, is critical for appropriate resource allocation for prevention planning and treatment and care services.
- The New York City Department of Health and Mental Hygiene (NYC DOHMH) maintains a population-based HIV/AIDS Registry and routinely ascertains local deaths based on matches with city death certificates.
- Persons in the Registry may die outside NYC if they are traveling or if they had moved outside NYC after being diagnosed and reported. To ascertain such deaths, sources other than city death certificates are needed.
- We evaluated the utility of two national death databases to ascertain deaths among PWHA in the Registry not known to be deceased. To our knowledge, this is the first comparison of these databases among PWHA.

METHODS

Population

- 32,970 persons with HIV/AIDS reported to the NYC HIV/AIDS Registry by 12/31/2005 not recently in care and not known to be deceased (Figure 1).

Data sources

- The **National Death Index (NDI)** is a database of death records from state vital statistics offices. It is maintained by the National Center for Health Statistics. The NDI contains names, SSNs, demographic data, and dates and cause(s) for all deaths occurring in the US and dependent territories reported to local jurisdictions. The NDI is updated annually and is considered the "gold standard" for mortality ascertainment.[1] There is a fee per case per year searched in the NDI.
- The **Social Security Administration's (SSA) Death Master File (SSDMF)** contains names, social security numbers (SSNs), and dates of birth for deaths reported to the SSA, usually in connection with a claim for death benefits. Cause of death is not available. The SSDMF is updated monthly and is available for a flat fee.
- The **NYC HIV/AIDS Registry** contains identifiers and clinical and demographic data for all persons diagnosed and reported with HIV/AIDS in NYC.

Matching process

- PWHA were independently matched to deaths in 2000-2004 reported in the SSDMF and NDI. Identifiers used for each match were based on availability of data in each source.
- NDI match
  - Initial matches were generated by NDI staff based on the following identifiers: first name, last name, date of birth, SSN, sex, race, place of residence, and place of birth. Probabilistic scores were assigned to each potential match.[2]
  - Exact matches were accepted after verification, along with those matching by all of the following identifiers: first name, last name, date of birth and SSN.
  - Other matches with a score over 28 were reviewed independently by two reviewers. Discordant pairs were resolved by a third reviewer.
- SSDMF match
  - Initial matches were performed using the following identifiers: first name, last name, date of birth and SSN. Cases that matched exactly by 1) first name, last name, and date of birth, or 2) SSN, were accepted after verification.
  - Other cases matching on some but not all fields were reviewed independently by two reviewers. Discordant pairs were resolved by a third reviewer.
- Matches were conducted using SAS 9.2 and Microsoft Access.

Analysis

- Number of deaths were ascertained overall and for each source.
- The % of deaths ascertained based on either source was determined overall and within groups (sex, age, race/ethnicity, HIV transmission category).
- To evaluate the usefulness of the SSDMF, the % of deaths uniquely ascertained by the SSDMF versus the NDI was calculated.
- The kappa statistic (κ) was calculated to assess agreement between sources. κ of 0.61-0.80 represented "substantial" agreement and 0.81-1.00 "almost perfect" agreement.[3]
- The number of persons reported to be living with HIV/AIDS (PLWHA) was determined before and after the match to assess % change in prevalence.
- Analyses were performed in SAS 9.2.

LIMITATIONS

- Results may not be generalizable to other disease populations and/or other jurisdictions.
- Different matching algorithms used for each source based on availability of identifiers may affect ascertainment.
  - No gold standard exists to confirm inexact matches.
- Variations in ascertainment by subgroup may reflect:
  - subgroup-specific differences in mortality (e.g., higher among injection drug users and older people)
  - differences in ease of matching personal identifiers (e.g., females harder to match due to surname changes)
  - mobility (e.g., non-native residents may be more likely to move prior to death)

RESULTS

Figure 1. Selection criteria and results of NDI and SSDMF matches with PWHA

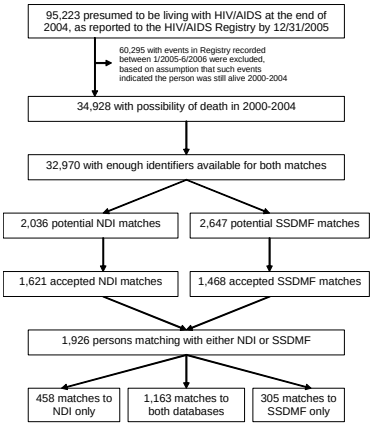


Table 1. Demographic characteristics of PWHA ascertained as dead in 2000-2004 by the NDI and SSDMF (N=1,926)

	Ascertained by NDI or SSDMF (N=1,926)		Ascertained by NDI (N=1,621)		Ascertained by SSMDF (N=1,468)	
	N	%	N	%	N	%
Sex						
Male	1,456	75.6	1,215	75.0	1,132	77.1
Female	470	24.4	406	25.0	336	22.9
Age at death						
<30	64	3.3	52	3.2	43	2.9
30-39	454	23.6	397	24.5	337	23.0
40-49	770	40.0	657	40.5	572	39.0
50-59	444	23.1	362	22.3	352	24.0
60+	194	10.1	153	9.4	164	11.2
Race/ethnicity						
Black	791	41.1	665	41.0	602	41.0
Hispanic	609	31.6	496	30.6	419	28.5
White	494	25.6	433	26.7	420	28.6
Other/Unknown*	32	1.7	27	1.7	27	1.8
HIV transmission risk						
Men who have sex with men	447	23.2	388	23.9	387	26.4
Injection drug use history	779	40.4	645	39.8	559	38.1
Heterosexual	276	14.3	227	14.0	216	14.7
Other risk	37	1.9	33	2.0	29	2.0
Unknown risk	387	20.1	328	20.2	277	18.9
State of death or receipt of benefits						
New York State	837	43.5	649	40.0	557	37.9
New York City only	*397*	*20.6*	*222*	*13.7*	*256*	*17.4*
Outside New York City	*440*	*22.8*	*427*	*26.3*	*301*	*20.5*
New Jersey	204	10.6	195	12.0	158	10.8
Florida	167	8.7	159	9.8	147	10.0
Puerto Rico	105	5.5	99	6.1	67	4.6
Pennsylvania	67	3.5	64	3.9	54	3.7
All others or unknown	546	28.3	455	28.1	485	33.0

*Other/Unknown includes Asian/Pacific Islander, Native American, and other and unknown race.

- **Of 32,970 PWHA evaluated, 1,926 (5.9%) were ascertained as dead, including 1,621 from the NDI and 1,488 from the SSDMF (Figure 1).**
- **Newly ascertained deaths were predominantly among males (75.6%), blacks and Hispanics (72.7%), and injection drug users (40.4%) (Table 1).**
- **Death were ascertained in 37 US states, Washington DC, and Puerto Rico, in addition to New York City and New York State.**
- **64.0% of deaths ascertained by the NDI were due to HIV; 32.5% were due to non-HIV-related causes.**
- **305 (15.8%) of all deaths were uniquely ascertained by the SSDMF. Such deaths were more likely to be among males (p=0.01), older people (p<0.001), and NYC residents at death (p<0.0001).**

Figure 2. % of cases with death ascertained

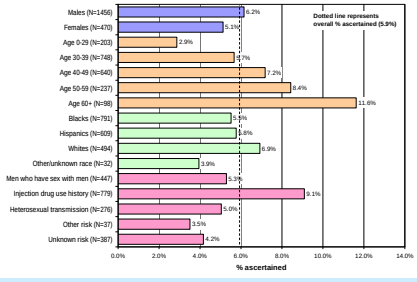
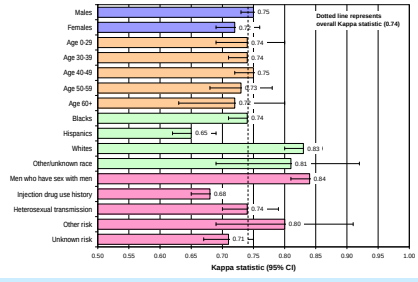


Figure 3. Agreement between NDI and SSDMF



- **Ascertainment was higher among injection drug users (9.1%), older people (11.6% in persons 60+), and whites (6.9%) (Figure 2).**
- **Kappa overall between NDI and SSDMF was 0.74 (95% CI 0.72-0.76) (Figure 3), representing substantial agreement between the two sources.**
- **Agreement was lower in certain subgroups (e.g., Hispanics) (Figure 3), suggesting more difficulty in matching identifiers in these groups.**

Table 2. HIV/AIDS prevalence estimates in New York City before and after matches

	Before match		After match		% change in N
	PLWHA (N)	Rate per 100 pop.	PLWHA (N)	Rate per 100 pop.	
Total	**95,223**	**1.19**	**93,297**	**1.17**	**-2.0%**
Sex*					
Male	66,112	1.74	64,657	1.70	-2.2%
Female	28,925	0.69	28,454	0.68	-1.6%
Age at end of 2004					
0-29	8,043	0.23	8,002	0.23	-0.5%
30-39	21,768	1.61	21,475	1.59	-1.3%
40-49	37,723	3.33	36,949	3.26	-2.1%
50-59	21,034	2.47	20,463	2.41	-2.7%
60+	6,655	0.53	6,408	0.51	-3.7%
Race/ethnicity					
Black	42,276	2.15	41,485	2.11	-1.9%
Hispanic	30,290	1.40	29,682	1.37	-2.0%
White	20,350	0.73	19,857	0.71	-2.4%
Other/Unknown**	2,307	0.21	2,273	0.21	-1.5%
HIV transmission risk					
Men who have sex with men	26,637	N/A	26,189	N/A	-1.7%
Injection drug use history	22,390	N/A	21,616	N/A	-3.5%
Heterosexual	17,635	N/A	17,359	N/A	-1.6%
Other risk	2,990	N/A	2,953	N/A	-1.2%
Unknown risk	25,571	N/A	25,180	N/A	-1.5%

*Excludes 186 persons with unknown sex. **Other/Unknown includes Asian/Pacific Islander, Native American, and other and unknown race.

HIV/AIDS prevalence estimates decreased by 2.0% (Table 2). This decrease was even greater within certain subgroups (e.g. injection drug users, 3.5%; persons 60+, 3.7%).

References: 1. Cowper 2002, Ann Epidemiol 12: 462-468. 2. Rogot 1986, J Chronic Dis 39: 719-734. 3. Landis 1977, Biometrics 33: 159-174.

CONCLUSIONS

- Matching with two national death databases identified 1,926 deaths in 5 years among PWHA reported in NYC who were previously believed to be alive.
  - Most deaths occurred outside New York City.
  - Prevalence estimates and rates were reduced by 0.5-3.7%, depending on the subgroup of interest.
- HIV/AIDS surveillance programs should conduct periodic death matches outside the local jurisdiction to improve accuracy of HIV/AIDS prevalence and mortality counts.
- Using both the NDI and the SSDMF increases the likelihood of finding a match and ascertaining death, because alternate forms of identifiers may exist in each database.
  - The NDI remains the primary source for identifying deaths occurring outside the local jurisdiction, due to the availability of accurate data (including cause of death) as recorded on death certificates.
  - The SSDMF is an inexpensive means to identify deaths beyond those found by the NDI, in order to maximize death ascertainment.

We would like to thank members of the HIV Epidemiology and Field Services Program, especially Marcia Bryan, Wanda Davis, Selam Seyoum, and Dipal Shah for review of matches, Sonny Ly and Julie Evans for data management, and Beth Begier for logistical guidance; and Richard Selik of the Division of HIV/AIDS Prevention, CDC, for providing access to the SSDMF data.

Presented at the 40th Annual Society for Epidemiologic Research Meeting, Boston, June 19-22, 2007.