

**Weekly**

June 13, 2008 / 57(23);631-634

# Electronic Record Linkage to Identify Deaths Among Persons with AIDS --- District of Columbia, 2000--2005

An estimated 1 million persons in the United States are living with human immunodeficiency virus (HIV)/acquired immunodeficiency syndrome (AIDS); approximately 500,000 persons with AIDS have died since 1981 (1,2). In 2005, the District of Columbia (DC) had an estimated adult AIDS prevalence rate of 2%, one of the highest AIDS prevalence rates in the United States (2). Accurate death ascertainment is an important part of HIV/AIDS surveillance. Manual methods can substantially underestimate deaths by missing death certificates that do not mention HIV infection or deaths of residents that occur in other states. CDC and the Council of State and Territorial Epidemiologists (CSTE) recommend performing electronic record linkages to ascertain deaths annually as part of routine HIV/AIDS surveillance activities (3). In 2007, to identify all deaths that occurred during 2000--2005 among persons with AIDS who resided or received their diagnosis in DC, the HIV/AIDS Administration of the DC Department of Health, with assistance from CDC, performed an electronic record linkage. This report summarizes the results of that linkage, which determined that 54% of deaths among persons with AIDS had not been reported previously to the DC HIV/AIDS Reporting System (HARS). The results indicated that electronic record linkage for death ascertainment is necessary to more accurately estimate the prevalence of persons living with HIV/AIDS.

HARS is a confidential, name-based reporting system developed by CDC to manage HIV/AIDS surveillance data. HARS contains vital status information but does not contain information on cause of death. Until November 2006, DC records in HARS were limited to AIDS patients because non-AIDS patients with HIV infection were not reported by name in DC. To perform the electronic record linkage, Link Plus, a free program developed at CDC (4), was used to link AIDS patients in the HARS data file to records in two other computer data files: 1) the DC Vital Records Division's electronic death certificate file (eDCF) and 2) the Social Security Administration's Death Master File (SSDMF). The eDCF includes all deaths that occur in DC, regardless of state of residence, and some deaths of DC residents that occur in Maryland or Virginia. The SSDMF contains information on all deaths reported to the Social Security Administration, regardless of state of residence or where the death occurred. The eDCF has information on causes of death, but the SSDMF does not.

Analysis was limited to deaths that occurred during 2000--2005. The variables used for record linkage were name, date of birth, Social Security number, and sex. Three linkages were performed (Figure). Linkage 1 and linkage 2 matched the HARS file to eDCF and SSDMF records, respectively, to identify deaths among persons listed in HARS with reported AIDS. HARS cases that were successfully linked to eDCF or SSDMF records were categorized by whether the death had been previously reported to HARS.

To identify potential new AIDS cases never previously reported to HARS, linkage 3 identified those death certificates within eDCF that indicated HIV infection as a cause of death but had not been linked to HARS

via linkage 1. To ensure that these HIV-specific death certificates did not match any previously reported AIDS cases in HARS, a manual search of HARS records was conducted for matches after not finding them by electronic linkage. The remaining nonmatching HIV-specific death certificates were then matched to associated medical records to confirm that decedents met the surveillance case definition for HIV infection (5,6). If medical records were unavailable to corroborate the death certificate information, HIV/AIDS remained unconfirmed for the decedent because the surveillance case definition for HIV infection cannot be met by a death certificate alone (3). Multiple logistic regression was performed, and adjusted odds ratios were calculated to examine factors independently associated with whether a death was previously unreported to HARS before the electronic record linkage.

Linkage 1 and linkage 2 identified 2,460 deaths that occurred during 2000--2005 among persons with AIDS. Of these deaths, 1,337 (54%) had not been reported previously to HARS (Table 1). Among these previously unreported deaths, 320 (24%) were linked only to eDCF, 577 (43%) were linked only to SSDMF, and 440 (33%) were linked to both (Table 1).

Cause of death information was available for 1,562 (63%) of the 2,460 deaths. The underlying cause of death was HIV infection in 1,056 deaths (68%) and other causes (not HIV infection) in 506 deaths (32%) (Table 2). Of those 506 deaths attributed to other underlying causes, 112 (22%) death certificates mentioned HIV infection as a contributing (but not underlying) cause of death. Leading causes of the 506 deaths included cardiovascular disease (112 [22%]); cancer (98 [19%]); infectious diseases other than HIV infection (72 [14%]); homicide, suicide, or unintentional injury (52 [10%]); and chronic liver disease (30 [6%]). In a multiple logistic regression analysis, previously unreported deaths were associated with an underlying cause of death other than HIV infection (adjusted odds ratio: 7.53) but not with race/ethnicity, transmission category, sex, or age (Table 2).

Electronic linkage 3 identified 216 death certificates in eDCF that mentioned HIV infection as a cause of death but did not electronically match that information with reported AIDS patients in the HARS data file and thus might represent previously unreported HIV/AIDS cases. Overall, 97 (45%) cases were confirmed as new HIV/AIDS cases based on information from medical records. Of the other potential cases, 69 (32%) were matched manually to HARS patients (and therefore represented previously reported cases missed by linkage 1); 29 (13%) had only death certificate evidence of HIV infection available and thus remained unconfirmed; and 21 (10%) had no mention of HIV on the printed death certificate or medical records and were assumed to be erroneous.

**Reported by:** *T Jolaosho, MHS, J Gauntt, MS, T West-Ojo, MPH, MSPH, HIV/AIDS Admin, District of Columbia Dept of Health; AD Castel, MD, Dept of Epidemiology and Biostatistics, George Washington Univ School of Public Health and Health Svcs, District of Columbia. RM Selik, MD, T Durant, PhD, Div of HIV/AIDS Prevention, National Center for HIV, Viral Hepatitis, STD, and TB Prevention; PJ Peters, MD, E Tai, MD, EIS officers, CDC.*

## Editorial Note:

This report provides the first comparison of electronic record linkage with manual methods of AIDS death ascertainment in the United States. More than half (54%) of deaths among AIDS patients during 2000--2005 in DC had not been reported to HARS and were discovered by electronic record linkage with eDCF and SSDMF. This suggests that electronic record linkage is essential for complete ascertainment of deaths among persons with HIV/AIDS and accurate estimations of HIV/AIDS prevalence.

Death ascertainment in DC has relied on vital records staff members manually reviewing death certificates and sending records that mention HIV to HIV/AIDS surveillance staff members, who then manually matched the death certificates to HARS. Because this manual method is dependent upon death certificates

mentioning HIV infection, deaths with non-HIV underlying causes were less likely to be reported as a death in a person with AIDS. Antiretroviral therapy has reduced the percentage of deaths attributed to HIV infection and, therefore, limited the effectiveness of a manual death ascertainment method (7,8). Electronic record linkage has the advantage of being able to find deaths from all causes among persons with HIV/AIDS because electronic linkage can efficiently process large numbers of death records without being limited to death records that mention HIV infection.

Manual review of death certificates for a state generally is limited to persons who died in that state. A more complete ascertainment of deaths requires electronic linkage to a national death data file, such as SSDMF or the National Death Index. In this study, SSDMF and eDCF provided complementary and independent death information, with most deaths linked to only one of these data files. More deaths linked to SSDMF than eDCF, underscoring the importance of linking to a national death data file. The National Death Index, accessible through CDC's National Center for Health Statistics (9), is a national death certificate data file that is not limited to decedents with a Social Security number and includes information on causes of death. The National Death Index could be used to help ascertain deaths among AIDS patients; however, the index is more expensive to researchers because of fees charged to remunerate the state vital records offices that compile the data.

The findings in this report are subject to at least one limitation. The DC findings might not be entirely generalizable to the 50 states because DC's close proximity to other states might increase the frequency of out-of-state deaths that are not reported to the DC Vital Records Division. The conditions that led to underestimation of deaths, however, including the frequency of deaths with causes other than HIV infection and the possibility that HIV-infected persons might die in another state, affect many areas of the United States.

Electronic linkage of the HIV/AIDS case registry with a state's death-certificate registry and with a national death registry such as SSDMF is a more efficient and thorough method to ascertain deaths among persons with HIV/AIDS than manual linkage limited to in-state death certificates that mention HIV. Improved death ascertainment can enable more accurate estimates of HIV/AIDS prevalence and a more effective allocation of HIV prevention and treatment resources. These findings support the CDC/CSTE recommendation to perform electronic record linkage to ascertain deaths annually as part of routine HIV/AIDS surveillance activities (3). Most state HIV/AIDS surveillance programs have followed this recommendation (CSTE, unpublished data, 2007), but a barrier in some areas has been the lack of personnel skilled at computer programming to perform these electronic linkages.

## Acknowledgments

This report is based, in part, on contributions by F Johnson-Clarke, PhD, Research and Analysis Div, and J Davidson-Randall, Vital Records Div, State Center for Health Statistics Admin, Dept of Health, District of Columbia.

## References

1. Glynn M, Rhodes P. Estimated HIV prevalence in the United States at the end of 2003. [Abstract] 2005 National HIV Prevention Conference, June 12--15, Atlanta, GA. Available at <http://www.aegis.com/conferences/NHIVPC/2005/T1-B1101.html>.
2. CDC. Cases of HIV infection and AIDS in the United States and dependent areas, 2005: HIV/AIDS surveillance report, 2005. Vol 17. Revised ed. Atlanta, GA: US Department of Health and Human Services, CDC; 2007. Available at <http://www.cdc.gov/hiv/topics/surveillance/resources/reports/2005report/default.htm>.
3. CDC and Council of State and Territorial Epidemiologists. Technical guidance for HIV/AIDS

surveillance programs, volume I: policies and procedures. Atlanta GA: US Department of Health and Human Services, CDC; 2005.

4. National Program of Cancer Registries. Link Plus. Atlanta, GA: US Department of Health and Human Services, CDC; 2007. Available at <http://www.cdc.gov/cancer/npcr/tools/registryplus/lp.htm>.
5. [CDC. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. MMWR 1992;41\(No. RR-17\):1--19.](#)
6. [CDC. Guidelines for national human immunodeficiency virus case surveillance, including monitoring for human immunodeficiency virus infection and acquired immunodeficiency syndrome. MMWR 1999; 48\(No. RR-18\).](#)
7. Mocroft A, Brettle R, Kirk O, et al. Changes in the cause of death among HIV positive subjects across Europe: results from the EuroSIDA study. *AIDS* 2002;16:1663--71.
8. Sackoff JE, Hanna DB, Pfeiffer MR, Torian LV. Causes of death among persons with AIDS in the era of highly active antiretroviral therapy: New York City. *Ann Intern Med* 2006;145:397--406.
9. CDC. National Death Index. Hyattsville, MD: US Department of Health and Human Services, CDC, National Center for Health Statistics. Available at <http://www.cdc.gov/nchs/ndi.htm>.

## Table 1

**TABLE 1. Number and percentage of deaths among persons with AIDS linked electronically to the electronic Death Certificate File (eDCF) and Social Security Death Master File (SSDMF), by reporting status — District of Columbia, 2000–2005**

Reporting status	Linked only to eDCF		Linked only to SSDMF		Linked to both eDCF and SSDMF		Known deaths not linked to either eDCF or SSDMF		Total
	No.	(%)	No.	(%)	No.	(%)	No.	(%)	
Deaths not previously reported to HARS*	320	(24)	577	(43)	440	(33)	—	—	1,337
Deaths previously reported to HARS	321	(29)	181	(16)	481	(43)	140	(12)	1,123
<b>Total deaths</b>	<b>641</b>	<b>(26)</b>	<b>758</b>	<b>(31)</b>	<b>921</b>	<b>(37)</b>	<b>140</b>	<b>(6)</b>	<b>2,460</b>

\* HIV/AIDS Reporting System.

[Return to top.](#)

## Table 2

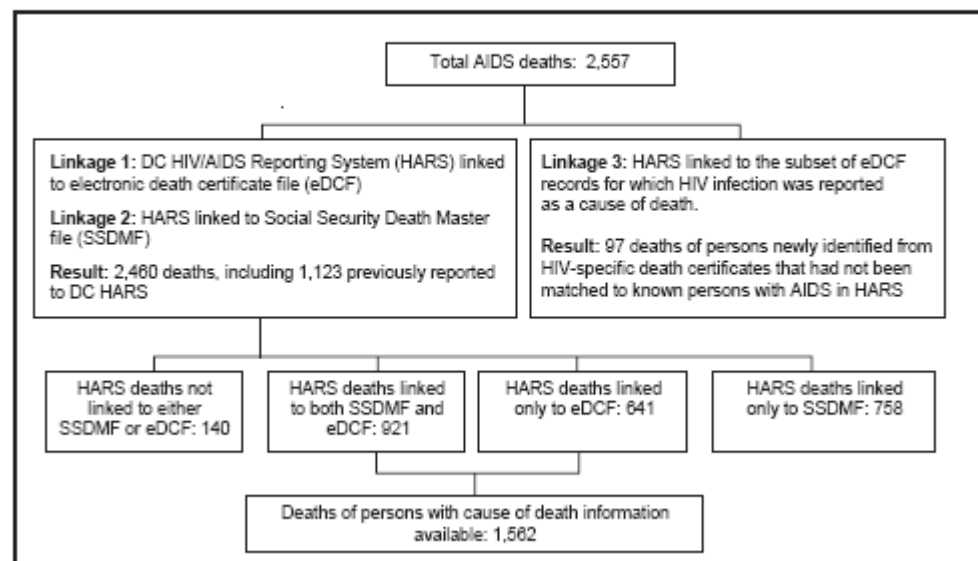
**TABLE 2. Number and percentage of deaths not previously reported to HARS,\* by selected characteristics — District of Columbia, 2000–2005**

Characteristic	Total deaths (n = 1,562 <sup>†</sup> )	Deaths not previously reported to HARS (n = 760)		Adjusted odds ratio	(95% CI) <sup>‡</sup>
		No.	(%)		
<b>Underlying cause of death<sup>¶</sup></b>					
HIV infection	1,056	367	(35)	Referent	—
Causes other than HIV infection	506	393	(78)	7.53	(5.80–9.79)
<b>Year of death</b>					
2000	272	99	(36)	Referent	—
2001	265	92	(35)	0.90	(0.60–1.34)
2002	276	119	(43)	1.37	(0.93–2.02)
2003	283	162	(57)	2.83	(1.93–4.14)
2004	268	165	(62)	3.65	(2.47–5.40)
2005	198	123	(62)	3.38	(2.21–5.17)
<b>Race/Ethnicity</b>					
Black, non-Hispanic	1,425	690	(48)	Referent	—
White, non-Hispanic	108	59	(55)	1.42	(0.88–2.29)
Hispanic/Other**/Not specified	29	11	(38)	0.73	(0.31–1.72)
<b>Transmission category</b>					
Illicit injection-drug use (IDU)	541	285	(53)	Referent	—
Men who have sex with men (MSM)	361	166	(46)	0.82	(0.58–1.15)
MSM and IDU	70	35	(50)	0.92	(0.51–1.64)
High-risk heterosexual contact <sup>††</sup>	297	132	(44)	0.76	(0.55–1.06)
No risk factor specified/Other <sup>§§</sup>	293	142	(48)	0.76	(0.54–1.05)
<b>Sex</b>					
Male	1,053	505	(48)	Referent	—
Female	509	255	(50)	1.18	(0.89–1.55)
<b>Age at death (yrs)</b>					
<40	408	182	(45)	Referent	—
40–49	653	309	(47)	0.87	(0.65–1.16)
50–59	385	203	(53)	0.99	(0.71–1.38)
≥60	116	66	(57)	0.78	(0.48–1.28)

\* HIV/AIDS Reporting System.

<sup>†</sup> Limited to the 1,562 deaths with underlying cause of death information from death certificate<sup>‡</sup> Confidence interval.<sup>¶</sup> Based on codes from the *International Classification of Diseases, Tenth Revision*.<sup>\*\*</sup> Non-Hispanic Asian, Pacific Islander, American Indian, or Alaska Native.<sup>††</sup> Sexual contact with a person known to be HIV-infected or at high risk for HIV infection (e.g., history of IDU or MSM).<sup>§§</sup> Includes mother-child transmission (n = 13) and transfusion (n = 3).[Return to top.](#)

## Figure

**FIGURE. Electronic linkages used to ascertain deaths among persons with AIDS — District of Columbia (DC), 2000–2005**[Return to top.](#)

Use of trade names and commercial sources is for identification only and does not imply endorsement by the U.S. Department of Health and Human Services.

References to non-CDC sites on the Internet are provided as a service to *MMWR* readers and do not constitute or imply endorsement of these organizations or their programs by CDC or the U.S. Department of Health and Human Services. CDC is not responsible for the content of pages found at these sites. URL addresses listed in *MMWR* were current as of the date of publication.

All *MMWR* HTML versions of articles are electronic conversions from typeset documents. This conversion might result in character translation or format errors in the HTML version. Users are referred to the electronic PDF version (<http://www.cdc.gov/mmwr>) and/or the original *MMWR* paper copy for printable versions of official text, figures, and tables. An original paper copy of this issue can be obtained from the Superintendent of Documents, U.S. Government Printing Office (GPO), Washington, DC 20402-9371; telephone: (202) 512-1800. Contact GPO for current prices.

\*\*Questions or messages regarding errors in formatting should be addressed to [mmwrq@cdc.gov](mailto:mmwrq@cdc.gov).

Date last reviewed: 6/11/2008

[HOME](#) | [ABOUT \*MMWR\*](#) | [MMWR SEARCH](#) | [DOWNLOADS](#) | [RSS](#) | [CONTACT](#)  
[POLICY](#) | [DISCLAIMER](#) | [ACCESSIBILITY](#)

GAFFB . MFAITMIB . DFADI E

**Morbidity and Mortality Weekly Report**  
Centers for Disease Control and Prevention  
1600 Clifton Rd, MailStop E-90, Atlanta, GA 30333,  
U.S.A



[Department of Health  
and Human Services](#)

