Federal Funding for Reporting Cases of HIV Infection in the United States, 2006

MATTHEW J. PAGE, MPP^{a,b,c}

KATHLEEN McDAVID HARRISON,
PhD, MPH^{a,c}

XIANGMING WEI, MS^{a,d}

H. IRENE HALL, PhD, MPH^a

SYNOPSIS

Objective. The Centers for Disease Control and Prevention (CDC) provides funding for human immunodeficiency virus (HIV) surveillance in 65 areas (states, cities, and U.S. dependent areas). We determined the amount of CDC funding per reported case of HIV infection and examined factors associated with differences in funding per reported case across areas.

Methods. We derived HIV data from the HIV/AIDS Reporting System (HARS) database. Budget numbers were based on award letters to health departments. We performed multivariate linear regression for all areas and for areas of low, moderate, and moderate-to-high morbidity.

Results. Mean funding per case reported was \$1,520, \$441, and \$411 in areas of low, moderate, and moderate-to-high morbidity, respectively. In low morbidity areas, funding per case decreased as log total cases increased ($p < 0.001$). For moderate and moderate-to-high morbidity areas, funding per case fell as log total cases increased ($p < 0.001$), but increased in accordance with an area's population ($p < 0.05$) and the proportion of that population residing in an urban setting ($p < 0.05$). The models for low, moderate, and moderate-to-high morbidity predicted funding per case as \$1,490, \$423, and \$390, respectively.

Conclusions. Economies of scale were evident. The amount of CDC core surveillance funding per case reported was significantly associated with the total number of cases in an area and, depending on morbidity, with total population and percentage of that population residing in an urban setting.

^aCenters for Disease Control and Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention/HIV Incidence and Case Surveillance Branch, Atlanta, GA

^bEmory University Rollins School of Public Health, Atlanta, GA

^cCurrent affiliation: Office of Health Equity, Centers for Disease Control and Prevention, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention/HIV Incidence and Case Surveillance Branch, Atlanta, GA

^dCurrent affiliation: Veterans Administration, Decatur, GA

Address correspondence to: Matthew J. Page, MPP, Office of Health Equity, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention/HIV Incidence and Case Surveillance Branch, Centers for Disease Control and Prevention, 1600 Clifton Rd. NE, MS E-07, Atlanta, GA 30333; tel. 404-639-0468; fax 404-639-8603; e-mail <MPage@cdc.gov>.

Since the beginning of the human immunodeficiency virus (HIV) epidemic almost 30 years ago, national HIV surveillance has evolved. In 1981, the Centers for Disease Control and Prevention (CDC) initiated population-based acquired immunodeficiency syndrome (AIDS) surveillance with a case definition based on opportunistic illnesses.¹ Since then, all states and dependent areas have conducted AIDS surveillance by using a standardized, confidential, name-based reporting system.² By 1985, serologic tests for the HIV antibody had become widely available, and several states had begun reporting HIV diagnoses. In late 1992, CDC published a revised definition for HIV and an expanded definition for AIDS.³ CDC published national guidelines for HIV case surveillance in 1999 and updated these with guidance in 2005.^{1,4} All 50 U.S. states, the District of Columbia, and five U.S. dependent areas had confidential name-based HIV infection reporting systems by April 2008 (Personal communication, Patricia Sweeney Hardy, CDC, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention/HIV Incidence and Case Surveillance Branch, April 2008).

CDC provides funding to states, cities, and U.S. dependent areas to monitor (1) the number and characteristics of people who have a diagnosis of HIV infection and (2) the trends in HIV- and AIDS-related morbidity and mortality.⁵ Surveillance data are used to elucidate changes in trends of HIV transmission; identify populations at risk; focus interventions and evaluate their effectiveness; allocate funds; and facilitate access to health, social, and prevention services.⁶ HIV surveillance comprises active and passive data collection. Regardless of how the information is received, all data are entered locally into the HIV/AIDS Reporting System (HARS). After personal identifying information has been removed, the data are sent to CDC.

For several years, the funding for core HIV surveillance in all areas has remained stable. However, the number of incident cases occurring in a given year has recently been estimated to be more than previously reported.⁷ In this context, it is increasingly important that limited funds be efficiently and accurately directed to populations at risk. The purpose of this study was to determine the amount of CDC core surveillance funding allocated per reported case of HIV and to examine factors associated with differences in funding per case reported during 2006.

METHODS

Areas included in analysis

HIV surveillance is currently funded in the 50 U.S. states and the District of Columbia (DC); six separately

funded cities (Chicago, Houston, Los Angeles, Philadelphia, New York City, and San Francisco); and five U.S. dependent areas (American Samoa, Guam, the Northern Mariana Islands, Puerto Rico, and the U.S. Virgin Islands).² Surveillance funding is also provided to the Federated States of Micronesia, the Marshall Islands, and Palau, which report only AIDS cases. The six cities receive funding separately from the states in which they are located because of high AIDS morbidity in the earlier years of the epidemic. In 2006, CDC accepted cases of HIV infection from 45 states and five dependent areas that conducted confidential name-based HIV infection reporting.²

We limited our analyses to the 50 states, DC, and the six separately funded cities. Cities were defined based on the area served by the relevant separately funded city health department. We included DC with the states rather than with the separately funded cities because DC is a stand-alone entity. We did not include U.S. dependent areas because of differences in reporting structures and funding mechanisms, and the lack of availability of specific variables.

HIV and AIDS cases

We used HIV data from the CDC HARS database.² We considered three issues when determining HIV diagnoses by reporting area. First, only data from states with mature (at least four years of data reporting to CDC), confidential, name-based HIV reporting systems are generally included in reports. However, because the maturity of the system was not our paramount concern, we included all 45 states that conducted confidential name-based HIV infection reporting during 2006. To estimate the number of HIV diagnoses in the other six areas, we calculated an AIDS:HIV ratio, which was based on the total number of AIDS and HIV diagnoses reported by those 45 states, and then divided the total number of AIDS cases in DC and each of the remaining five states (Hawaii, Maryland, Massachusetts, Montana, and Vermont) by this ratio.

The second factor was our need to use the reporting city or state rather than area of residence for people diagnosed with HIV infection (area of residence is routinely used in CDC publications). Therefore, the numbers used in our analyses may be slightly different from those presented in CDC's annual HIV/AIDS Surveillance Report. However, although the reporting state is a routinely collected variable, health departments are not required to name the reporting city when they submit data to HARS. This difference in reporting requirements was of particular concern when trying to determine the number of cases reported by the five states with separately funded cities. However, in each of the states with separately funded cities, the percentage

for which the HIV reporting city was available was higher than the national average (71.4%): 100.0% in Illinois, New York, and Texas; 99.5% in California; and 85.3% in Pennsylvania. Communication with the state health department allowed us to determine the reporting city for the remaining 14.7% of Pennsylvania HIV diagnoses.

The third factor was progression from HIV infection to AIDS (i.e., HIV diagnoses reported in 2006 that were later reported as AIDS). The total number of HIV diagnoses reported by any given area comprises two groups—HIV only (i.e., reported as HIV in 2006, but not later reported as AIDS) and HIV progressed to AIDS (i.e., reported as HIV in 2006, but later reported as AIDS, regardless of whether the AIDS report was made in 2006 or later or whether the report was made by the original area or a different area). In 2006, 8.1% (4,517/55,498) of diagnoses reported as HIV were later reported as AIDS by the same area. The cases that progressed to AIDS in 2006 are included in the AIDS numbers for 2006. However, because our concern was the cost of reporting cases of HIV and AIDS, “double counting” was not an issue. In other words, these cases were included twice because they were reported twice—once for the HIV report and once for the AIDS report; therefore, the cost of reporting was incurred twice.

Area HIV surveillance budgets

For our analyses, we used the total 2006 core surveillance budgets based on the “Notice of Grant Award” letters sent to each city and state health department. Core surveillance budgets comprise direct costs, including salary/wages, fringe benefits, travel, equipment, supplies, and contractual costs, as well as indirect costs. When more than one award was allocated or an award was rescinded, the net final amount was used in the analysis.

Calculation of outcome variable (CDC core HIV surveillance funding per case)

CDC core HIV surveillance funding per case reported was calculated by dividing the 2006 total core surveillance budget (direct and indirect costs) in a specific area by the combined number of HIV and AIDS diagnoses reported in 2006. Because funding is for both HIV and AIDS surveillance, we combined HIV and AIDS cases when calculating the average funding per case reported.

Explanatory variables

Using data from HARS, we included the number of HIV and AIDS cases as well as the percentage of cases

in each transmission category—male-to-male sexual contact, injection drug use (IDU), high-risk heterosexual contact, male-to-male sexual contact and IDU, and no identified risk factor.² The percentage of cases in each transmission category was included because surveillance funding decisions may be based, to some degree, on the prevalence of certain risk groups. Population-based data for 2006 were obtained from the U.S. Census Bureau and other federal sources. In our initial analyses, we included the total population in an area, percentage black, percentage male, median age,⁸ and percentage foreign-born.^{9,10} We also considered the region of the country in which the area is located, proportion of the population living in an urban setting,¹¹ and area in square miles of the funded city or state.^{12,13} The percentage of foreign-born people is for 2005,^{9,10} while the percentage who were urban residents is based on estimates for 2000.¹¹

To approximate overall costs in a particular area, we included median annual household income (values for the 50 states and DC were derived from U.S. Census estimates for 2006; estimates of city income were based on 2003 U.S. Census estimates that were updated to 2006 using the seasonally adjusted U.S. city-average consumer price index for all items).^{14–16} Two variables were included as approximations of labor costs. First was the annual median salary for an epidemiologist, which was derived from Bureau of Labor Statistics occupational wage categories.¹⁷ For the states for which the salary of an epidemiologist was unavailable ($n=20$), we imputed the salary by multiplying each state’s percentage of the median all-occupation salary by the median salary for the states that did report the salary for an epidemiologist. We followed a similar procedure for the two separately funded cities for which a median salary for an epidemiologist was not reported. The second variable was the cost of living, estimated via the fair-market rent for a two-bedroom apartment.¹⁸

Finally, believing that the quality of reporting might affect the cost of reporting, we developed a proxy for the quality of surveillance data. This proxy was the number of years since an area initiated HIV reporting of any kind (code- or name-based). This variable was calculated by subtracting the month and year in which such reporting was initiated from December 31, 2006. All separately funded cities except Philadelphia were assumed to have initiated HIV reporting in the same month and year as the state in which they are located.²

Analyses

The only funds considered in the analyses were those provided by CDC to the states and cities to conduct

core surveillance. No state- or city-provided funds or other costs incurred by CDC were considered in the analyses. All costs are reported in 2006 U.S. dollars.

We used SAS[®] version 9.1 to perform all analyses¹⁹ and calculated initial descriptive statistics. Next, we assessed variables for normality and linearity as related to the outcome variable (funding per case reported). In developing the models, we first assessed correlation of all potential explanatory variables with the outcome variable (cut point = 0.30). Then we evaluated collinearity using condition indices (cut point = 0.30) and variance decomposition proportions (cut point = 0.50). We used manual backward elimination (alpha = 0.05) to assess potential effect modifiers (all two-way interaction terms) and potential confounders and to fit a parsimonious linear regression model based on the full model remaining after assessment of collinearity.²⁰ Backward elimination was conducted for each of four data stratifications: (1) all areas, (2) areas with low morbidity, (3) areas with moderate morbidity, and (4) areas with moderate-to-high morbidity (Figure 1).

All data were collected as part of routine HIV surveillance as mandated by state or local laws or regulations. CDC determined that this project was not a research activity and, thus, did not require review by an Institutional Review Board.

RESULTS

Descriptive statistics

The mean number of reported HIV cases in 2006 was 1,396 for the states (plus DC) and 4,268 for the six cities. Total cases ranged from eight to 9,371 (median = 766) for the states and from 2,039 to 8,522 for the cities (median = 3,336). Average total CDC funding for core surveillance in 2006 was \$486,603 for the states and \$1,229,985 for the cities (data not shown).

Funding per case reported was \$1,000 or more for 13 states and less than \$100 for one of the states. For the states, the mean funding per case reported was \$879. For the separately funded cities, the mean funding per case reported was \$315. The funding per case reported did not exceed \$1,000 or cost less than \$100 for any of the cities. Of the 13 areas whose funding per case was more than \$1,000, 12 were areas of low morbidity and one was an area of moderate morbidity. The one state whose funding was less than \$100 was an area of high morbidity (data not shown).

As shown in Table 1, in low morbidity areas, the mean number of reported HIV cases in 2006 was 239, ranging from eight to 934 cases (median = 171). For moderate and high morbidity areas, the mean total cases were 1,815 and 5,594, respectively. The mean

funding per case of HIV reported for low, moderate, and high morbidity areas was \$1,520, \$441, and \$287, respectively. In moderate-to-high morbidity areas, the mean of total cases was 2,550 and the mean funding per case reported was \$411.

Model development

We initially developed four models: (1) an overall model encompassing all areas, (2) a low morbidity model, (3) a moderate morbidity model, and (4) a high morbidity model. The high morbidity model would not run because it contained more potential explanatory variables than observations ($n=7$). Therefore, we combined the moderate and high morbidity models. First, we evaluated all potential explanatory variables for normality as well as for the degree of linearity relative to funding per case reported. No gross deviations from

Figure 1. AIDS morbidity level in U.S. cities and states, 2006

Low morbidity areas <225 AIDS cases ^a (n=21)	Moderate morbidity areas 225–1,500 AIDS cases ^a (n=29)	High morbidity areas >1,500 AIDS cases ^a (n=7)
Alaska	Alabama	California
Delaware	Arizona	Florida
Hawaii	Arkansas	Georgia
Idaho	Chicago ^b	Maryland
Iowa	Colorado	New York
Kansas	Connecticut	New York City ^b
Kentucky	District of Columbia	Texas
Maine	Houston ^b	
Minnesota	Illinois	
Montana	Indiana	
Nebraska	Los Angeles ^b	
New Hampshire	Louisiana	
New Mexico	Massachusetts	
North Dakota	Michigan	
Oklahoma	Mississippi	
Rhode Island	Missouri	
South Dakota	Nevada	
Utah	New Jersey	
Vermont	North Carolina	
West Virginia	Ohio	
Wyoming	Oregon	
	Pennsylvania	
	Philadelphia ^b	
	San Francisco ^b	
	South Carolina	
	Tennessee	
	Virginia	
	Washington	
	Wisconsin	

^aReported cases

^bSeparately funded city

AIDS = acquired immunodeficiency syndrome

Table 1. Reported HIV cases, total CDC funding for core HIV surveillance, and funding per case reported, by AIDS morbidity level, U.S., 2006

<i>AIDS morbidity level</i>	<i>Mean (SD)</i>	<i>Median</i>	<i>Range</i>
Low morbidity (n=21)			
Total HIV cases reported	239 (237)	171	8–934
Area HIV surveillance budget (in 2006 dollars)	\$152,329 (\$84,136)	\$133,070	\$61,003–\$359,280
Funding per HIV case reported (in 2006 dollars)	\$1,520 (\$1,716)	\$1,030	\$139–\$7,916
Moderate morbidity (n=29)			
Total HIV cases reported	1,815 (1,302)	1,548	309–6,082
Area HIV surveillance budget (in 2006 dollars)	\$630,466 (\$380,547)	\$522,515	\$216,490–\$2,093,595
Funding per HIV case reported (in 2006 dollars)	\$441 (\$234)	\$366	\$107–\$1,076
Moderate-to-high morbidity (n=36)			
Total HIV cases reported	2,550 (2,156)	1,647	309–9,371
Area HIV surveillance budget (in 2006 dollars)	\$805,493 (\$597,416)	\$579,870	\$216,490–\$2,959,772
Funding per HIV case reported (in 2006 dollars)	\$411 (\$226)	\$355	\$97–\$1,076

HIV = human immunodeficiency virus

CDC = Centers for Disease Control and Prevention

AIDS = acquired immunodeficiency syndrome

SD = standard deviation

normality were seen for any of the potential explanatory variables. We transformed the outcome variable from funding per case reported to the natural log of funding per case reported. We also transformed cases by taking the natural log of each observation. This process yielded relationships between the outcome variable and potential explanatory variables that were closer to linear. For all further analyses, we used log funding per case reported as the outcome variable and log cases as one of the explanatory variables.

We evaluated the degree of correlation for each potential explanatory variable with the log funding per case reported. This led us to eliminate 12 of the 18 potential explanatory variables. The potential explanatory variables eliminated based on evaluation of correlation were percentage of cases reporting male-to-male sexual contact, IDU, high-risk heterosexual contact, male-to-male sexual contact and IDU, and no identified risk factor; median age; percentage foreign-born; region of the country; area in square miles; median annual household income; median salary for an epidemiologist; and fair-market rent for a two-bedroom apartment.

Next, based on our collinearity assessment, we eliminated seven of the 15 possible two-way interaction terms and one of the potential explanatory variables (percentage male). The full starting model included log total cases; total population; percentage black; percentage urban; and years since HIV surveillance was initiated, as well as eight interaction terms: (log cases * total population; log cases * years since HIV surveillance initiated; total population * percentage urban; total population * percentage black; total population

* years since HIV surveillance initiated; percentage urban * percentage black; percentage urban * years since HIV surveillance initiated; and percentage black * years since HIV surveillance initiated).

Parsimonious models

We performed backward elimination by first removing the most insignificant interaction term regardless of whether its *p*-value was the highest among all the insignificant terms. Next, to verify the models obtained through our backward elimination procedure, we performed a second backward elimination procedure, removing all insignificant terms in order (highest to lowest *p*-value) as well as a forward selection procedure. None of the interaction terms in any of the models was statistically significant.

The overall parsimonious model showed that log funding per case reported decreased as log total cases increased ($p<0.001$) (Table 2). Alternatively, the log funding per case reported increased as total population ($p<0.05$) and the percentage of the population residing in an urban setting ($p<0.05$) increased. In all areas, total cases reported ranged from eight to 9,371 with a mean of 1,698 (standard deviation [SD] = 2,047) and a median of 1,032. The average total population was 5,252,605 (SD=4,952,225) with a median of 4,206,074 and range of 515,005 to 25,765,427. Overall, the proportion of the population living in an urban setting ranged from 0.38 to 1.00 with a mean of 0.75 (SD=0.17) and a median of 0.74.

In the low morbidity model, log funding per case reported decreased as log total cases ($p<0.001$) increased. No other variables were significant for the

low morbidity model. In these areas, the mean number of cases reported was 239 (SD=237).

For moderate morbidity areas, the parsimonious model showed that the log funding per case reported decreased as log total cases increased ($p<0.001$). The log funding per case reported increased in accordance with an area's population ($p<0.05$) and the proportion of an area's population residing in an urban setting ($p<0.05$). In moderate morbidity areas, the mean number of cases reported was 1,815 (SD=1,302). The mean total population and proportion in an urban setting were 5,634,263 (SD=2,927,750) and 0.79 (SD=0.16), respectively.

After we included the seven high morbidity areas in the moderate morbidity model, the parsimonious model again showed that the log funding per case reported decreased as log total cases increased ($p<0.001$) and increased as the area's population ($p<0.05$) and percentage urban ($p<0.05$) increased. In these areas, the mean number of cases reported was 2,550 (SD=2,156). While the mean total population in these areas was 7,253,544 (SD=5,207,814), the mean proportion of the population residing in an urban area was 0.81 (SD=0.15).

We used the four parsimonious models to estimate the predicted funding per case reported for each level of morbidity (Figure 2). In the overall parsimonious model, containing log cases, total population, and percentage urban, the predicted funding per case reported ranged from \$173 to \$5,572, with a mean predicted value of \$763 (Figure 2, panel A). In the low morbidity model, which was based on log cases only, the predicted funding per case reported ranged from \$256

to \$7,053 (Figure 2, panel B), with a mean of \$1,490. For the moderate and moderate-to-high morbidity models, both containing log cases, total population, and percentage urban, the predicted funding per case reported ranged from \$163 to \$1,139 and from \$167 to \$1,077, respectively (Figure 2, panels C and D). The mean predicted funding per case reported was \$423 for the moderate morbidity model and \$390 for the moderate-to-high morbidity model.

DISCUSSION

In all four models, an increase in log total cases was statistically significantly associated with a decrease in the log funding per case reported. This economies-of-scale phenomenon has also been seen in an analysis of the cost of reporting cancer cases.²¹ As in the cancer funding analysis, the economies-of-scale relationship in HIV surveillance may be largely due to fixed costs incurred in all areas. That is, every city or state reporting HIV cases to CDC encounters certain baseline administrative and logistical costs that are similar, regardless of morbidity level. This finding suggests that surveillance programs whose output potential is limited might achieve gains in operating efficiency by sharing fixed costs and other essential resources.

In the overall, moderate, and moderate-to-high morbidity models, funding per case reported increased as the total population in an area increased. However, the coefficients for total population were quite small in all three models (Table 2). For these three models, funding per case reported also increased as the percentage of the population residing in an urban area increased.

Table 2. Linear regression model of log average CDC funding per HIV case reported, U.S., 2006

Covariates	Coefficients (p-value)			
	All areas (n=57)	Low morbidity areas (n=21)	Moderate morbidity areas (n=29)	Moderate-to-high morbidity areas ^a (n=36)
Intercept	9.19	0.31	10.25	9.80
Log total cases	-0.63 ^b	-0.70 ^b	-0.77 ^b	-0.71 ^c
Total population	4.46E-08 ^c	NI	5.20E-08 ^c	3.71E-08 ^b
Percentage residing in urban area	1.29 ^c	NI	1.31 ^c	1.40 ^b
Adjusted R ²	0.80	0.83	0.61	0.55

^aThe high morbidity model would not run because it contained more potential explanatory variables than observations. Therefore, the moderate and high morbidity models were merged to form a moderate-to-high morbidity model.

^bStatistically significant at $p<0.001$

^cStatistically significant at $p<0.05$

CDC = Centers for Disease Control and Prevention

HIV = human immunodeficiency virus

NI = not included

Figure 2. Mean Centers for Disease Control and Prevention funding per human immunodeficiency virus case reported, U.S., 2006

A. All areas

NOTE: For the model covering all areas, the observed log funding per case reported ranged from 4.6 to 9.0 and the predicted log funding ranged from 5.2 to 8.6. While the observed funding per case for the overall model ranged from \$97 to \$7,916, the predicted funding per case ranged from \$173 to \$5,572.

B. Low morbidity areas

NOTE: For the model covering low morbidity areas, the observed log funding per case reported ranged from 4.9 to 9.0 and the predicted log funding ranged from 5.5 to 8.9. While the observed funding per case for the low morbidity model ranged from \$139 to \$7,916, the predicted funding per case ranged from \$256 to \$7,053. The low morbidity states included Alaska, Delaware, Hawaii, Idaho, Iowa, Kansas, Kentucky, Maine, Minnesota, Montana, Nebraska, New Hampshire, New Mexico, North Dakota, Oklahoma, Rhode Island, South Dakota, Utah, Vermont, West Virginia, and Wyoming.

Figure 2 (continued). Mean CDC funding per HIV case reported, U.S., 2006**C. Moderate morbidity areas**

NOTE: For the model covering moderate morbidity areas, the observed log funding per case reported ranged from 4.7 to 7.0 and the predicted log funding ranged from 5.1 to 7.0. While the observed funding per case for the low morbidity model ranged from \$107 to \$1,076, the predicted funding per case ranged from \$163 to \$1,139. The moderate morbidity areas included Alabama, Arizona, Arkansas, Chicago (Illinois), Colorado, Connecticut, District of Columbia, Houston (Texas), Illinois (excluding Chicago), Indiana, Los Angeles, Louisiana, Massachusetts, Michigan, Mississippi, Missouri, Nevada, New Jersey, North Carolina, Ohio, Oregon, Pennsylvania (excluding Philadelphia), Philadelphia (Pennsylvania), San Francisco (California), South Carolina, Tennessee, Virginia, Washington, and Wisconsin.

D. Moderate-to-high morbidity areas

NOTE: For the model covering moderate-to-high morbidity areas, the observed log funding per case reported ranged from 4.6 to 7.0 and the predicted log funding ranged from 5.1 to 7.0. While the observed funding per case for the low morbidity model ranged from \$97 to \$1,076, the predicted funding per case ranged from \$167 to \$1,077. The high morbidity model would not run because it contained more potential explanatory variables than observations. Therefore, the moderate and high morbidity models were merged to form a moderate-to-high morbidity model. The moderate morbidity areas included Alabama, Arizona, Arkansas, Chicago (Illinois), Colorado, Connecticut, District of Columbia, Houston (Texas), Illinois (excluding Chicago), Indiana, Los Angeles, Louisiana, Massachusetts, Michigan, Mississippi, Missouri, Nevada, New Jersey, North Carolina, Ohio, Oregon, Pennsylvania (excluding Philadelphia), Philadelphia (Pennsylvania), San Francisco (California), South Carolina, Tennessee, Virginia, Washington, and Wisconsin. The high morbidity areas included California (excluding Los Angeles and San Francisco), Florida, Georgia, Maryland, New York State (excluding New York City), New York City (New York), and Texas (excluding Houston).

Both of these associations make sense in that all costs, including those related to HIV surveillance, tend to be higher in large population centers.

Rather than showing the true cost of reporting a case, our analyses show which dependent variables are associated with CDC core HIV surveillance funding. In other words, the four models developed as part of this analysis show how CDC core HIV surveillance funds are allocated by case of HIV reported. Ideally, the funding per case reported would be very close to the true cost of reporting a case. Some areas, however, may be “doing more with less.” In such areas, the true cost of reporting a case would be higher than the funding received per case. Ideally, funding decisions are made in such a way that fewer states, given available funding, need to do more with less. Our analyses have elucidated a descriptive model that shows the factors related to current funding levels based on fixed total amounts of funding available.

A logical next analytical step would be the development of a model to determine how funding decisions could be based on the true cost of reporting a case. Assessing the true cost of reporting a case would involve looking at several factors, including the costs of individual surveillance activities such as setting up electronic reporting and maintaining it on an ongoing basis. Such an assessment would also have to address the issue of electronic laboratory reporting to find cases. Specifically, are laboratory reports of new cases obtained electronically or manually? If electronic, how much work is involved with—and how much cost incurred by—programming to make the process smooth and routine?

In addition, the analysis does not include costs incurred by areas related to obtaining updates on the progression from HIV to AIDS. For instance, are updates to AIDS obtained through staff review of medical records in search of low CD4 counts and/or AIDS indicator diseases, or are updates to AIDS obtained from laboratory reports of low CD4 counts? If the latter is the case, costs would vary depending on what proportion of the reporting is electronic as well as the method for entering these laboratory data (manual or batch uploading) into HARS.

Limitations

Our analyses were subject to some limitations. First, we imputed HIV diagnoses for five states and DC on the basis of the AIDS:HIV ratio in the remaining 45 states. This method yielded a relatively accurate estimate, although not the exact number of HIV diagnoses in any of the five states or DC.

Further, imputing the number of HIV diagnoses may

have led to an underestimation of the funding per case reported for those areas. As mentioned previously, cases reported as HIV in 2006 and reported as AIDS later in 2006 were included twice in the analysis. One issue with this methodology is that cases of HIV that progressed to AIDS are categorized by reporting area. Therefore, if one area reported a case as HIV and another area reported that case as AIDS, both cases were attributed to the AIDS-reporting area. However, there is no reason to assume that areas experienced differing levels of in-migration of people reported with HIV infection in the origination area in 2006 and reported with AIDS in the destination area later that year.

We included years since the initiation of any type of HIV reporting as a proxy for data quality. This decision was based on the assumption that it takes time for the quality of data from a reporting system to improve. In this context, quality relates to the accuracy of case information as well as completeness of case ascertainment (i.e., all cases diagnosed are reported). If the accuracy of information was low, then funding per case reported may have been overestimated. If more true cases were actually reported (i.e., improved completeness of case ascertainment), funding per case reported would have decreased. In other words, more cases would be reported with the same amount of total surveillance funding. This would increase the denominator (total cases reported) while the numerator (total surveillance funding) would remain constant. While the completeness of AIDS reporting has been reported as 85% or better,^{22,23} completeness of HIV reporting has been estimated to be 76% six months after diagnosis and 81% 12 months after diagnosis.²⁴

Given recent emphasis on increasing the transition from paper to electronic medical records, a main issue is the existence and ability of the program to obtain access to electronic medical records. This access will have a large impact on completeness and timeliness of reporting as well as costs, including those costs covered by areas' supplemental monies. Areas' supplemental monies (in addition to CDC surveillance funds) were not included in the analysis.

In addition, analyses did not include CDC's surveillance-related costs (e.g., review and processing of grant applications, and database management). Also, because some components of CDC surveillance funding (e.g., equipment and travel) may fluctuate from year to year, 2006 funding in some areas might not be representative of average annual funding in those areas. Finally, the seven areas transitioning from code-based to name-based reporting of HIV cases in 2006 may have reported more cases in that year than they will in the future once all prevalent HIV cases have been

reported. This overreporting reduced the amount of funding per case reported in these seven areas.

CONCLUSIONS

The amount of CDC core surveillance funding per case of HIV reported is significantly associated with the total number of cases in an area and, depending on the area's level of morbidity, with total population and the percentage of that population residing in an urban setting. Our descriptive cost models elucidate the factors associated with current surveillance funding and provide a strong foundation for the development of models to determine how funding decisions could be made in the future.

The authors thank the public health advisors for reviewing the budget information for accuracy; the states and cities for collecting and reporting case data; Ram Shrestha for carefully reviewing the manuscript; and Marie Morgan for editing and reviewing the manuscript.

The findings and conclusions in this article are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

REFERENCES

1. Fleming PL, Ward JW, Janssen RS, De Cock KM. Guidelines for national human immunodeficiency virus case surveillance, including monitoring for human immunodeficiency virus infection and acquired immunodeficiency syndrome. *MMWR Recomm Rep* 1999;48(RR-13):1-28.
2. Centers for Disease Control and Prevention (US). HIV/AIDS surveillance report, 2006. Vol. 18. 2008 [cited 2010 Apr 28]. Available from: URL: <http://www.cdc.gov/hiv/topics/surveillance/resources/reports/2006report/default.htm>
3. Castro KG, Ward JW, Slutsker L, Buehler JW, Jaffe HW, Berkelman RL. 1993 revised classification system for HIV infection and expanded surveillance case definition for AIDS among adolescents and adults. *MMWR Recomm Rep* 1992;41(RR-17):1-18.
4. Hall HI, Mokotoff ED; Advisory Group for Technical Guidance on HIV/AIDS Surveillance. Setting standards and an evaluation framework for human immunodeficiency virus/acquired immunodeficiency syndrome surveillance. *J Public Health Manag Pract* 2007;13:519-23.
5. Centers for Disease Control and Prevention (US). HIV prevention strategic plan: extended through 2010. October 2007 [cited 2009 Mar 25]. Available from: URL: <http://www.cdc.gov/hiv/resources/reports/psp>
6. McDavid K, McKenna MT. HIV/AIDS risk factor ascertainment: a critical challenge. *AIDS Patient Care STDS* 2006;20:285-92.
7. Hall HI, Song R, Rhodes P, Prejean J, An Q, Lee LM, et al.; HIV Incidence Surveillance Group. Estimation of HIV incidence in the United States. *JAMA* 2008;300:520-9.
8. National Center for Health Statistics (US). Postcensal estimates of the resident population of the United States for July 1, 2000–July 1, 2006, by year, county, age, bridged race, Hispanic origin, and sex (vintage 2006) [cited 2010 Apr 28]. Available from: URL: <http://www.cdc.gov/nchs/about/major/dvs/popbridge/popbridge.htm>
9. Census Bureau (US). Statistical abstract of the United States, 2008: Table 40. Native and foreign-born population by state: 2005. Washington: Census Bureau; 2007. Also available from: URL: <http://www.census.gov/prod/2007pubs/08abstract/pop.pdf> [cited 2009 Mar 30].
10. Census Bureau (US). Statistical abstract of the United States, 2008: Table 41. Nativity and place of birth of resident population—25 largest cities: 2005. Washington: Census Bureau; 2007. Also available from: URL: <http://www.census.gov/prod/2007pubs/08abstract/pop.pdf> [cited 2009 Mar 30].
11. Census Bureau (US). Statistical abstract of the United States, 2008: Table 29. Urban and rural population by state: 1990 and 2000. Washington: Census Bureau; 2007. Also available from: URL: <http://www.census.gov/prod/2007pubs/08abstract/pop.pdf> [cited 2009 Mar 30].
12. Census Bureau (US). Population size, density and distribution for states and Puerto Rico. 2005 [cited 2010 Apr 28]. Available from: URL: http://www.census.gov/popest/gallery/maps/popsiz_2006.html
13. Census Bureau (US). Population of the 100 largest cities and other urban places in the United States: 1790 to 1990 [cited 2010 Apr 28]. Available from: URL: <http://www.census.gov/population/www/documentation/twps0027/twps0027.html>
14. Census Bureau (US). 2006 American Community Survey. R1901. Median household income (in 2006 inflation-adjusted dollars): 2006 universe: households [cited 2009 Mar 25]. Available from: URL: http://factfinder.census.gov/servlet/GRTTable?_bm=y&-geo_id=01000US&-box_head_nbr=R1901&-ds_name=ACS_2006_EST_G00_&-lang=en&-redoLog=false&-format=US-30&-mt_name=ACS_2006_EST_G00_R0501_US30&-CONTEXT=grt
15. Census Bureau (US). American Community Survey. Median household income (in 2003 inflation-adjusted dollars) [cited 2009 Mar 25]. Available from: URL: <http://www.census.gov/acs/www/Products/Ranking/2003/R07T160.htm>
16. Department of Labor (US), Bureau of Labor Statistics. Consumer price indexes. Customized tables for 2006 and 2003 seasonally adjusted U.S. city-average indexes for all items [cited 2010 Apr 28]. Available from: URL: <http://www.bls.gov/CPI>
17. Department of Labor (US), Bureau of Labor Statistics. Occupational employment statistics. Annual median wage for "epidemiologist"—191041. Washington: Department of Labor; May 2006.
18. Department of Housing and Urban Development (US). Fair market rents. 2006 final data—county level data file [cited 2009 Mar 25]. Available from: URL: <http://www.huduser.org/datasets/fmr.html>
19. SAS Institute, Inc. SAS®: Version 9.1 for Windows. Cary (NC): SAS Institute, Inc.; 2004.
20. Kleinbaum DG, Kupper LL, Muller KE, Nizam A. Applied regression analysis and multivariable methods. 3rd ed. Pacific Grove (CA): Duxbury Press; 1998.
21. Weir HK, Berg GD, Mansley EC, Belloni KA. The National Program of Cancer Registries: explaining state variations in average cost per case reported. *Prev Chronic Dis* 2005;2:A10. Also available from: URL: http://www.cdc.gov/pcd/issues/2005/jul/04_0124.htm [cited 2009 Feb 22].
22. Hall HI, Li J, McKenna MT. HIV in predominantly rural areas of the United States. *J Rural Health* 2005;21:245-53.
23. Kleven RM, Fleming PL, Li J, Gaines CG, Gallagher K, Schwarcz S, et al. The completeness, validity, and timeliness of AIDS surveillance data. *Ann Epidemiol* 2001;11:443-9.
24. Hall HI, Song R, Gerstle JE 3rd, Lee LM; HIV/AIDS Reporting System Evaluation Group. Assessing the completeness of reporting of human immunodeficiency virus diagnoses in 2002–2003: capture-recapture methods. *Am J Epidemiol* 2006;164:391-7.