

## SURVEILLANCE OF HIV INFECTION

This report includes data from case reports from 56 areas (all 50 states, the District of Columbia, American Samoa, Guam, the Northern Mariana Islands, Puerto Rico, and the U.S. Virgin Islands) that had laws or regulations requiring confidential reporting by name for adults, adolescents, and children with confirmed diagnoses of HIV infection (including AIDS) as of December 31, 2009. After the removal of personal identifying information, data from these reports were submitted to CDC. The implementation of HIV infection reporting has differed from state to state. Before 1991, the surveillance of HIV infection was not standardized, and reporting was based primarily on passive surveillance. The information on many of the cases reported before 1991 is not complete. Since then, CDC has assisted states in conducting active surveillance of HIV infection by the use of standardized report forms and software. The Republic of Palau received authorization to begin reporting HIV surveillance data to CDC in 2010; these data will be included in future reports.

Data on diagnoses of HIV infection should be interpreted with caution. HIV surveillance reports may not be representative of all persons infected with HIV because not all infected persons have been tested. Many HIV-reporting states offer anonymous HIV testing; the results of anonymous tests are not reported to the confidential name-based HIV registries of state and local health departments. Therefore, reports of confidential test results may not represent all persons who tested positive for HIV infection. Furthermore, many factors, including the extent to which testing is routinely offered to specific groups and the availability of, and access to, medical care and testing services, may influence testing patterns. These data provide a minimum estimate of persons known to be HIV infected. In addition, because surveillance practices differ, the reporting and the updating of a person's clinical and vital status differ among states. The completeness of reporting of HIV infection is estimated at more than 80% [1].

Because states initiated confidential name-based HIV infection reporting on different dates, the length of time that reporting has been in place influences the number of diagnoses of HIV infection reported to

CDC. For example, data presented for the first year a state initiated reporting may include cases diagnosed during only part of the year (i.e., case reporting may have begun after January).

For this report, we used the 2008 revised HIV case definition to classify HIV infection among adults and adolescents and among children [2]. The revised definition increases the specificity of the case definition by requiring laboratory-confirmed evidence of HIV infection. Increased specificity results in more accurate data on the number of diagnoses of HIV infection; these data, in turn, can be used to refine public health policies and determine the most appropriate use of HIV resources.

## TABULATION AND PRESENTATION OF DATA

The data in this report are provisional. This report includes information received by CDC through June 30, 2010. The data are organized into 4 sections.

- Section 1 (Tables 1a/b–10a/b): numbers (statistically adjusted and unadjusted) and rates (adjusted) of diagnoses of HIV infection and AIDS
- Section 2 (Tables 11a/b–14a/b): numbers (statistically adjusted and unadjusted) and rates (adjusted) of deaths and proportions (unadjusted) of survival among persons with a diagnosis of HIV infection or AIDS
- Section 3 (Tables 15a/b–18a/b): numbers (statistically adjusted and unadjusted) and rates (adjusted) of persons living with a diagnosis of HIV infection or AIDS
- Section 4 (Tables 19–24): numbers (statistically adjusted and unadjusted) and rates (adjusted) of diagnoses of HIV infection and AIDS, and persons living with a diagnosis of HIV infection or AIDS, presented by state and metropolitan statistical area (MSA)

For the assessment of trends in diagnoses, deaths, or prevalence, it is preferable to use statistically adjusted (estimated) data to eliminate artifacts of reporting in the surveillance system. Rates based on estimated numbers less than 12 should be interpreted with caution because the estimates have relative standard errors greater than 30% and are considered unreliable.

## **Areas with Mature HIV Infection Reporting Systems**

An area's confidential name-based HIV infection reporting is considered mature after 4 years—long enough for the calculation of reporting-delay estimates and the determination of reliable trends [3]. To estimate the number of diagnoses of HIV infection, we used data from 45 areas (i.e., 40 states and 5 U.S. dependent areas) with laws or regulations requiring confidential name-based HIV infection reporting since at least January 2006. The 45 areas comprise 40 states (Alabama, Alaska, Arizona, Arkansas, Colorado, Connecticut, Florida, Georgia, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Michigan, Minnesota, Mississippi, Missouri, Nebraska, New Hampshire, New Jersey, New Mexico, New York, Nevada, North Carolina, North Dakota, Ohio, Oklahoma, Pennsylvania, South Carolina, South Dakota, Tennessee, Texas, Utah, Virginia, West Virginia, Wisconsin, and Wyoming) and 5 U.S. dependent areas (American Samoa, Guam, the Northern Mariana Islands, Puerto Rico, and the U.S. Virgin Islands). For Tables 1a/b, 3a/b, 5a/b, 8, 10a/b, 19, and 23, we used data from these 45 areas to tabulate numbers (unadjusted), estimated numbers, and estimated rates of diagnoses of HIV infection. These data were also used to tabulate the numbers and the estimated numbers and rates of persons living with HIV infection (Tables 15a/b, 17a/b, 21, and 23).

## **Areas with Confidential Name-Based HIV Infection Reporting**

All 56 areas had implemented confidential name-based HIV infection reporting as of April 2008 and were included in tabulations of numbers (unadjusted) of diagnoses of HIV infection by state and MSA. For Tables 19 and 21, we used data from all 50 states, the District of Columbia, and 5 U.S. dependent areas to calculate numbers (unadjusted) of diagnoses of HIV infection. For Table 23, we used data from all 50 states, the District of Columbia, and Puerto Rico to calculate numbers (unadjusted) of diagnoses of HIV infection.

## **Deaths**

Persons reported to the national HIV surveillance system are assumed alive unless their deaths have been reported to CDC. Death data (Tables 11a/b–12a/b) include all deaths of persons with a diagnosis of HIV

infection or AIDS regardless of the cause of death. Because of delays in the reporting of deaths, 3 years (2006–2008) of death data are displayed. The exclusion of data from the most recent year allowed at least 18 months for deaths to be reported to CDC. The estimated numbers and rates of deaths resulted from statistical adjustment for delays in reporting.

## **Survival Analyses**

For the survival analyses presented in Section 2 (Tables 13a/b–14a/b), we used the Kaplan-Meier method to estimate the probability of survival for 12, 24, and 36 months for persons with a diagnosis of HIV infection or AIDS whose case data were reported by June 30, 2010. Tables were limited to persons whose diagnosis was made during 2001–2005 to allow at least 3 years from the time of diagnosis to deaths occurring through December 31, 2008. Data for each HIV reporting area were included in the survival tables beginning with the first full calendar year after implementation of code-based or name-based HIV infection reporting (e.g., Connecticut implemented code-based HIV infection reporting in January 2002, so in this report, the Connecticut data in the survival tables are for HIV infections diagnosed during 2002–2005).

## **Persons Living with a Diagnosis of HIV Infection or AIDS**

Because of delays in the reporting of deaths, 3 years of prevalence data are displayed. The data reflect the persons living with a diagnosis of HIV infection or AIDS at the end of each year during 2006–2008. The exclusion of data from the most recent year allowed at least 18 months for deaths to be reported and for these deaths to be factored into calculations of prevalence (persons living with HIV infection or AIDS).

## **Age**

The designation “adults and adolescents” refers to persons aged 13 years and older; the designation “children” refers to persons aged less than 13 years. For presentations of data on persons living with HIV infection or AIDS (Tables 15a/b–18a/b, 21, and 22), the age-group assignment or the age designation is based on the person's age as of December 31, 2008. In Tables 11a/b and 12a/b, which concern deaths of persons with a diagnosis of HIV infection or AIDS, age-group assignment is determined by the person's

age at the time of death. For all other tables, the age designation (for example, “adults and adolescents”) or the specific age-group assignment (for example, 20–24 years) is based on the person’s age at the time of diagnosis (HIV infection or AIDS).

### Race and Ethnicity

In the *Federal Register* for October 30, 1997 [4], the Office of Management and Budget (OMB) announced the Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. Implementation by January 1, 2003, was mandated. At a minimum, data on the following racial categories should be collected:

- American Indian or Alaska Native
- Asian
- black or African American
- Native Hawaiian or other Pacific Islander
- white

Additionally, systems must be able to retain information when multiple racial categories are reported. In addition to data on race, data on 2 categories of ethnicity should be collected:

- Hispanic or Latino
- not Hispanic or Latino

The Asian or Pacific Islander category displayed in previous surveillance reports has been split into 2 categories: (1) Asian and (2) Native Hawaiian or other Pacific Islander. The Asian category (in tables where footnoted) includes the cases in Asians/Pacific Islanders (referred to as legacy cases) that were reported before the implementation of the new racial categories in 2003 (e.g., cases of HIV infection that were diagnosed and reported to CDC before 2003 but that progressed to AIDS after 2003) and a small percentage of cases that were reported after 2003 but that were reported according to the old racial category (Asian/Pacific Islander). In tables of diagnoses of HIV infection during 2006–2009, the Asian category does not include Asian/Pacific Islander cases because these cases were diagnosed after 2003 and were reported to CDC in accordance with OMB’s Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity [4].

This report also presents data for persons who reported multiple racial categories. Persons whose race was unknown are included in the total numbers in

Tables 11a/b–18a/b, 21, and 22, and in the cumulative totals of Tables 2a/b and 6a/b. The number of persons reported in each race category may include persons whose ethnicity was not reported. In this report, persons categorized by race were not Hispanic or Latino.

### Geographic Designations

The areas of residence included in the report are defined as follows:

**Northeast:** Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont

**Midwest:** Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin

**South:** Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia

**West:** Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming

**U.S. dependent areas:** American Samoa, Guam, the Northern Mariana Islands, Puerto Rico, and the U.S. Virgin Islands

### Metropolitan Statistical Areas

In the *Federal Register* for December 27, 2000, the OMB published revised standards for defining MSAs in federal statistical activities [5]. These standards, which provided for the identification of MSAs in the United States and Puerto Rico, replaced the 1990 standards. The adoption of the new standards was effective as of December 27, 2000. On June 6, 2003, the OMB announced new MSA definitions based on the new standards and Census 2000 data [6]. Tables 23 (HIV diagnosis data) and 24 (AIDS data) present numbers and rates of diagnoses, by MSA, for areas with populations of 500,000 or more. The MSAs listed in Tables 23 and 24 are defined according to the OMB’s most recent update (November 2008) of statistical areas [7].

### Transmission Categories

*Transmission category* is the term for the classification of cases that summarizes a person’s possible HIV risk factors; the summary classification results from selecting, from the presumed hierarchical order of probability, the one risk factor most likely to have

been responsible for transmission. For surveillance purposes, a diagnosis of HIV infection or AIDS is counted only once in the hierarchy of transmission categories. Persons with more than one reported risk factor for HIV infection are classified in the transmission category listed first in the hierarchy. The exception is men who report sexual contact with other men *and* injection drug use; this group makes up a separate transmission category.

Persons whose transmission category is classified as male-to-male sexual contact include men who report sexual contact with other men (i.e., homosexual contact) and men who report sexual contact with both men and women (i.e., bisexual contact). Persons whose transmission category is classified as heterosexual contact are persons who report specific heterosexual contact with a person known to have, or to be at high risk for, HIV infection (e.g., an injection drug user). The term *high-risk* is not included in the transmission category label for *heterosexual contact* in the tables because heterosexual contact itself is the risk factor most likely to have been responsible for transmission; however, the table footnote regarding this category clarifies how the data are defined: “heterosexual contact with a person known to have, or to be at high risk for, HIV infection.”

Adults and adolescents born in, or who had sex with someone born in, a country where heterosexual transmission was believed to be the predominant mode of HIV transmission (formerly classified as Pattern II countries by the World Health Organization) are no longer classified as having heterosexually acquired HIV infection unless they meet the criteria stated in the preceding paragraph. Similar to other cases in persons who were reported without information about a behavioral or a transfusion risk factor for HIV infection, these cases are classified (in the absence of other risk factor information that would classify them in another transmission category) as “no risk factor reported or identified” [8]. Cases in children whose mother was born in, or whose mother had sex with someone born in, a Pattern II country are now classified (in the absence of other risk factor information that would classify them in another transmission category) as “mother with documented HIV infection, a risk factor for HIV infection, or HIV infection without a specified risk factor.”

Cases in persons with no reported risk factor for HIV infection listed in the hierarchy of transmission categories are classified as “no risk factor reported or

identified.” No identified risk factor (NIR) cases include cases that have been followed up by local health department officials; cases in persons whose risk-factor information is missing because they died, declined to be interviewed, or were lost to follow-up; and cases in persons who were interviewed or for whom other follow-up information was available but for whom no risk factor was identified.

Because a substantial proportion of cases of HIV infection and AIDS are reported to CDC without an identified risk factor, multiple imputation is used to assign a risk factor for these cases [9]. Multiple imputation is a statistical approach in which each missing risk factor is replaced with a set of plausible values that represent the uncertainty about the true, but missing, value [10]. The plausible values are analyzed by using standard procedures, and the results from these analyses are then combined to produce the final results. In this report, multiple imputation has been used in tables displaying estimated values, by transmission category, for adults and adolescents, but not in tables displaying data for children (because the number of cases in children is small, missing risk factors were not imputed).

## Reporting Delays

Reporting delays (time between diagnosis or death and the reporting of diagnosis or death to CDC) may differ among demographic and geographic categories; for some, delays in reporting have been as long as several years. The statistical adjustment of the data on diagnoses and deaths is based on estimates of reporting-delay distributions, which are calculated by using a modified semiparametric life-table statistical procedure. This procedure takes into account differences in reporting delays due to sex, race/ethnicity, and HIV transmission categories; reporting city, state, or territory; geographic region; the size of the MSA; and the type of facility where the diagnosis was made or death occurred. In addition, this method accounts for changes in the patterns of reporting delays over time as well as reporting delays of more than 5 years [3].

## Rates

Rates per 100,000 population were calculated for the numbers of diagnoses of HIV infection and AIDS, the deaths of persons with a diagnosis of HIV infection or AIDS, and persons living with a diagnosis of HIV infection or AIDS. The population denominators used to compute these rates for the 50 states, the District of

Columbia, and Puerto Rico were based on the official postcensus estimates for 2009 from the U.S. Census Bureau [11]. The population denominators for American Samoa, Guam, the Northern Mariana Islands, and the U.S. Virgin Islands were based on estimates and projections from the U.S. Census Bureau's International Data Base [12]. Each rate was calculated by dividing the estimated total number of diagnoses (or deaths, or persons living with a diagnosis of HIV infection or AIDS) for the calendar year by the population for that calendar year and then multiplying the number by 100,000. The denominators used for calculating age-, sex-, and race/ethnicity-specific rates were computed by applying the 2009 postcensus estimates for age, sex, and race/ethnicity for the 50 states and the District of Columbia [11]. The same method was used to calculate the denominators for Puerto Rico, with the exception of race/ethnicity estimates; these data are not available for Puerto Rico. For the other 4 U.S. dependent areas, estimates from the U.S. Census Bureau's International Data Base were used for age- and sex-specific population denominators [12].

CDC currently does not provide subpopulation rates by race/ethnicity for the 5 U.S. dependent areas because the U.S. Census Bureau does not collect information from all U.S. dependent areas. Similarly, rates for transmission categories are not provided because of the absence of denominator data (i.e., the denominator data used in this report come from the U.S. Census Bureau, but the U.S. Census Bureau does not collect data on transmission categories).

## REFERENCES

- Hall HI, Song R, Gerstle JE III, Lee LM; on behalf of the 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;164:391–397.
- CDC. Revised surveillance case definitions for HIV infection among adults, adolescents, and children aged <18 months and for HIV infection and AIDS among children aged 18 months to <13 years—United States, 2008. *MMWR* 2008;57(RR-10):1–12.
- Song R, Hall HI, Frey R. Uncertainties associated with incidence estimates of HIV/AIDS diagnoses adjusted for reporting delay and risk redistribution. *Stat Med* 2005;24:453–464.
- Office of Management and Budget. Revisions to the standards for the classification of federal data on race and ethnicity. *Federal Register* 1997;62:58781–58790. [http://www.whitehouse.gov/omb/fedreg\\_1997standards](http://www.whitehouse.gov/omb/fedreg_1997standards). Accessed November 2, 2010.
- Office of Management and Budget. Standards for defining metropolitan and micropolitan statistical areas. *Federal Register* 2000;65(249):82228–82238. <http://www.whitehouse.gov/sites/default/files/omb/fedreg/metroareas122700.pdf>. Accessed November 2, 2010.
- Office of Management and Budget. Revised definitions of metropolitan statistical areas, new definitions of micropolitan statistical areas and combined statistical areas, and guidance on uses of the statistical definitions of these areas. OMB Bulletin 03-04. <http://www.whitehouse.gov/omb/bulletins/b03-04.html>. Published June 6, 2003. Accessed November 2, 2010.
- Office of Management and Budget. Update of statistical area definitions and guidance on their uses. OMB Bulletin 08-01. <http://www.whitehouse.gov/omb/bulletins/fy2008/b08-01.pdf>. Published November 20, 2007. Accessed November 2, 2010.
- CDC. Current trends: heterosexually acquired AIDS—United States, 1993. *MMWR* 1994;43:155–160.
- McDavid Harrison K, Kajese T, Hall HI, Song R. Risk factor redistribution of the national HIV/AIDS surveillance data: an alternative approach. *Public Health Rep* 2008;123(5):618–627.
- Rubin, DB. *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons Inc; 1987.
- U.S. Census Bureau. Population estimates [entire data set]. July 1, 2010. <http://www.census.gov/popest/estimates.php>. Published June 2010. Accessed June 10, 2010.
- U.S. Census Bureau. International Data Base. <http://www.census.gov/ipc/www/idb/>. Published June 2010. Accessed September 9, 2010.