Classification of Transmission Risk in the National HIV/AIDS Surveillance System

Lisa M. Lee, PhD[a]
Matthew T. McKenna, MD[a]
Robert S. Janssen, MD[a]

SYNOPSIS

Risk behavior information is essential for allocating resources and developing effective HIV prevention strategies. Over time, transmission risk information on HIV/AIDS cases has been less likely to be reported to the national surveillance system. The Centers for Disease Control and Prevention (CDC) invited approximately 30 experts in HIV/AIDS and behavioral research from state and local health departments, academia, community-based organizations, and the CDC to participate in a consultation in December 2001 to generate ideas on how best to deal with the lack of risk data. The group was charged with providing recommendations on methods for classifying and reporting risk information and for identifying methods and sources for improving ascertainment of transmission risk behaviors for individuals infected with HIV. This report describes the recommendations and the effects of implementing such recommended procedures on the national HIV/AIDS surveillance database.

[a]Centers for Disease Control and Prevention, Division of HIV/AIDS Prevention, Atlanta GA

Address correspondence to: Lisa M. Lee, PhD, 1600 Clifton Rd. NE, MS E-47, Atlanta GA 30333; tel. 404-639-2050; e-mail <LMLee@cdc.gov>.

HIV and AIDS case surveillance in the United States is defined as the ongoing and systematic collection, analysis, interpretation, dissemination, and evaluation of population-based information about people infected with HIV or diagnosed with AIDS. AIDS has been a reportable condition since the early 1980s in all 50 states, the District of Columbia, and U.S. dependencies and possessions in free association with the U.S. These areas report AIDS cases to the Centers for Disease Control and Prevention (CDC) using a uniform case definition and report form that includes information on demographic characteristics, clinical and laboratory information, and transmission risk behavior history. HIV case reporting has been part of several states' comprehensive HIV/AIDS surveillance systems since 1985. As of May 2003, 39 reporting areas have implemented name-based HIV case reporting as an extension of their AIDS case reporting system. Sixteen areas have implemented HIV case reporting using alternate or coded patient identifiers, and two reporting areas do not require any type of HIV case reporting. The CDC began integrating HIV case reporting from states with name-based systems into the AIDS reporting system in 1994.

The primary goals of the national HIV/AIDS surveillance system are to monitor the epidemic by counting cases and estimating incidence and prevalence, and to elucidate the epidemiology of HIV by answering questions related to person, place, and time. This information is then disseminated to people and organizations with decision-making and resource allocation responsibilities. For example, in 2002, AIDS data from the national HIV/AIDS surveillance system were used as part of an allocation formula to fund over $1.9 billion in treatment and care programs that are administered through the federal Health Resources and Services Administration (HRSA). Ultimately, the purpose of these data collection efforts is to guide public health programs and practice to prevent new HIV infections and prevent morbidity and mortality in those already infected.

HIV prevention services and care resources are often allocated based on relative disease burden in transmission risk groups. Data for transmission risk history are submitted to the HIV/AIDS reporting system on completed case report forms from health care providers or trained health department personnel abstracting in- and outpatient medical records. At some sites, supplemental surveillance activities are used to augment standard medical record reviews.' It is rare that risk behaviors are collected as part of routine disease surveillance systems for other diseases, and not all health care providers ascertain this information about patients or about their patients' partners. Ascertainment of patients' partners' risk behaviors is especially challenging, but is necessary for classification into the heterosexual contact (HC) risk category, which is defined as heterosexual contact with a person at risk for or infected with HIV.

At the start of the AIDS epidemic, when efforts were underway to describe the epidemiology of an unknown pathogen, most medical records contained information on transmission-related behaviors. With the current widespread availability of testing and greater awareness of transmission risks, less transmission behavior information is available through medical records. After a consistent and low percentage of AIDS cases with no risk reported (NRR) through the 1980s, there has been an increase in the proportion of NRR cases, with a high of 33% in 2000. Further, due to the greater number of cases to be investigated, there have been fewer NRR cases classified into a known risk category. Of AIDS NRR cases reported in 1992, 64% were eventually classified, whereas of NRR cases reported in 2000, only 37% were classified by the end of 2002. The effect of these two factors has been an increase in the accumulation of cases without transmission risk behavior information in the national system.

In the face of a growing number of cases initially reported without risk, the CDC developed statistical methods to mitigate bias introduced into the data from missing risk behavior. Using historical patterns of classification of NRR cases, the proportion of NRR cases classified into a known risk category is calculated. These fractions are then used to distribute current NRR cases into risk categories, stratifying by sex, race, and region. While risk is not imputed for individual cases, the distribution of cases into risk categories is adjusted. This method requires two major assumptions. First, it assumes that the distribution of true risk among NRR cases is homogeneous over time, which is unlikely to be true since changes in high-risk behaviors in some groups has changed the risk distribution over the past decade. Second, the method assumes that NRR cases that are classified are representative of all NRR cases. This is less likely to be true now than in the past, as an increased caseload has prohibited systematic follow-up of all NRR cases, leaving a convenience sample of cases that are chosen for investigation or resolved through passive ascertainment from later reports.

Since the assumptions of the risk redistribution have become increasingly untenable, CDC invited approximately 30 experts in HIV/AIDS and behavioral research from state and local health departments, academia, community-based organizations, and the CDC to par-

ticipate in a consultation in December 2001 to generate ideas on how best to deal with the lack of risk data. The group was charged with providing recommendations on methods for classifying and reporting risk information and for identifying methods and sources for improving ascertainment of transmission risk behaviors for individuals infected with HIV.

The consultation group recommended two changes in how the CDC presents HIV/AIDS transmission risk data. The first recommendation was to adopt a probable heterosexual contact (PHC) category to the list of current modes of exposure. PHC would be defined as having one or more partners of the opposite sex and denial of all other potential modes of exposure (e.g., male-to-male sex [MSM], injection drug use [IDU], blood/blood products). Denial of other modes must be stated (i.e., cannot be "unknown"). The PHC category would be in addition to the current HC category and would differ from it by not requiring identification of the partners' risk for HIV transmission. Based on transmission probabilities and difficulty ruling out other risk behaviors among men, including the stigma associated with homosexual activity, it was thought that PHC would be more relevant for female cases than male cases.

The second recommended change in how the CDC presents risk data was to stop using the hierarchy of most probable mode of transmission (Table 1) and report mutually exclusive categories of all risk behaviors. The group recommended the mutually exclusive categories shown in Table 2.

Concurrently, in an attempt to improve the characterization of HIV transmission in Michigan, Schmidt et al. described the use of supplemental HIV/AIDS surveillance projects to augment data reported to the HIV/AIDS surveillance system. Using additional medical record review data from a longitudinal medical record review project and a validation study of HIV transmission risk factors, the authors examined whether cases had additional risk factors that were not captured by using the CDC hierarchy of risk categories. Schmidt et al. proposed two changes to the current presentation of risk information. First, they recommended adding an IDU/HC category to capture the possible sexual transmission among IDU cases. Second, they recommended that the CDC add a "presumed heterosexual contact" category for women that includes cases that "after a thorough risk evaluation, cannot be classified as IDU or HC in the risk hierarchy, but who have been sexually active with one or more men." They suggested that if national data support it, the category should be applied to men as well.

We used available data in the national HIV/AIDS Reporting System (HARS) to assess the usefulness of these recommended changes for reporting of risk information for describing national transmission risk behaviors.

METHODS

We applied the two data presentation recommendations to the existing national HIV/AIDS surveillance data. To classify cases with insufficient risk information into the probable heterosexual category, we used sex of the reported case and the available ancillary behavior risk information. Male cases without an assigned risk were classified as PHC if their case report form stated "no" for IDU, sex with male, receipt of blood products and tissue, worked in an health care or laboratory setting, and HC with someone known to

Table 1. CDC hierarchy of the most probable mode of HIV transmission for adult/adolescent cases, 1993 to present

Most probable mode of HIV transmission
• Male-to-male sexual contact (MSM)
• Injection drug use (IDU)
• MSM and IDU
• Hemophilia
• Heterosexual contact with a person at risk for or infected with HIV
• Transfusion, transplant
• Confirmed other
• Risk not specified

Table 2. Mutually exclusive HIV transmission risk classification categories recommended by expert consultants, December 2001

Mutually exclusive HIV transmission risk classification categories
• Male-to-male sexual contact (MSM) only
• MSM and injection drug use (IDU)
• MSM and heterosexual contact with a person at risk for or infected with HIV (HC)
• MSM and IDU and HC
• IDU and HC
• HC
• Probable heterosexual contact (PHC)[a]
• Other (blood, tissue)

[a]PHC defined as having one or more partners of the opposite sex and denial of all other potential modes of exposure (e.g., MSM, IDU, blood/blood products).

be at risk for or infected with HIV, and stated "yes" for sex with female. Female cases without an assigned risk were classified as PHC if their case report form said "no" for IDU, receipt of blood products and tissue, worked in an health care or laboratory setting, and HC with someone known to be at risk for or infected with HIV, and stated "yes" for sex with male. The distribution of risk was compared when using risk as reported, risk as currently redistributed, and risk as reported including the PHC category. Because of concern that the PHC category would contain individuals (especially men) who truly have other, more probable risk behaviors, we also examined the proportion of individuals with known MSM or IDU risk that could be classified as PHC if they had denied MSM and/or IDU.

To present risk behavior data without the hierarchy of the most probable transmission mode, we developed mutually exclusive categories using all reported risk behaviors. Upon closer examination of the consultation recommendation, we reconstructed the categories to include several that had not been suggested. Our final list contained the 17 categories shown in Table 3. The distribution of risk behaviors was compared using risk as reported, risk as currently redistributed, and risk in mutually exclusive categories.

Analyses were performed using national HIV/AIDS surveillance data and were based on cases diagnosed in 1999, 2000, and 2001 and reported through September 2002. We used AIDS case data from all reporting areas and HIV case data from 29 states that have had name-based HIV reporting since 1997 and are considered systems from which new HIV diagnoses for this period can be examined. (The states are: Alabama, Arizona, Arkansas, Colorado, Florida, Idaho, Indiana, Iowa, Louisiana, Michigan, Minnesota, Missouri, Mississippi, Nevada, New Jersey, North Carolina, Nebraska, New Mexico, North Dakota, Ohio, Oklahoma, South Carolina, South Dakota, Tennessee, Utah, Virginia, West Virginia, Wisconsin, and Wyoming.)

Table 3. Modified mutually exclusive HIV transmission risk categories used in analysis of existing HIV/AIDS surveillance data

Modified mutually exclusive HIV transmission risk categories
• Male-to-male sexual contact (MSM) only
• Injection drug use (IDU) only
• Heterosexual contact with a person at risk for or infected with HIV (HC) only
• Presumed heterosexual contact (PHC)[a]
• Other (blood, tissue, healthcare worker) only
• MSM and IDU
• MSM and HC
• MSM and other
• IDU and HC
• IDU and other
• HC and other
• MSM and IDU and HC
• MSM and IDU and other
• IDU and HC and other
• MSM and IDU and HC and other
• No risk reported (at initial case report)
• No identified risk (after investigation)

[a]PHC defined as having one or more partners of the opposite sex and denial of all other potential modes of exposure (e.g., MSM, IDU, blood/blood products).

RESULTS

Presumed heterosexual contact

Table 4 compares the distribution of transmission risk at initial report, after standard redistribution, and using the PHC category for AIDS cases. AIDS case data from all reporting areas resulted in 122,794 cases reported from 1999 through 2001. Among 91,448 men reported with AIDS, 18,994 (21%) were initially reported without transmission risk. When PHC was computed from extant data, 1,240 (1.4%) men were considered PHC, reducing the proportion with no transmission risk to 19.4%. When risk was redistributed using the standard method, the proportion with no transmission risk was reduced to 0.2%. The proportion in all other categories increased as those without risk were distributed according to historical patterns. Among 31,346 women reported with AIDS, 9,633 (30.7%) were initially reported without transmission risk. When PHC was computed from extant data, 642 (2.1%) women were considered PHC, reducing the proportion with no transmission risk to 28.7%. When risk was redistributed using the standard method, the proportion with no transmission risk was reduced to 0.43%. The proportion in all other categories increased as those without risk were distributed according to historical patterns.

Table 5 compares the distribution of transmission risk at initial report, after standard redistribution, and using the PHC category for initial HIV diagnoses. Initial HIV diagnosis data from the 29 reporting areas resulted in 44,078 cases that had not progressed to AIDS reported from 1999 through 2001. Among 29,410 men reported with an initial HIV diagnosis, 9,532 (32.4%) were initially reported without transmission risk. Using the standard method, when risk was redis-

Table 4. Percentage distribution of HIV transmission risk as reported, after standard redistribution, with the addition of presumed heterosexual contact (PHC), and using mutually exclusive risk categories, AIDS diagnoses[a]—1999–2001, U.S., *n*=122,794

	Risk as reported, using current hierarchy		*Risk redistributed, using current statistical redistribution method*		*PHC, using recommended definition*		*Mutually exclusive categories, using recommended categories*	
	Male	*Female*	*Male*	*Female*	*Male*	*Female*	*Male*	*Female*
MSM	44.0	—	53.2	—	44.0	—	40.9	—
IDU	19.0	24.3	24.9	33.0	19.0	24.3	14.7	15.5
MSM/IDU	5.3	—	6.5	—	5.3	—	4.2	—
HC	10.2	43.5	14.1	64.2	10.2	43.5	10.1	42.9
Other	0.7	1.5	1.1	2.4	0.7	1.5	0.7	1.4
Undetermined	20.8	30.7	0.2	0.4	19.4	28.7	19.4	28.7
PHC					1.4	2.1	1.4	2.1
MSM/HC							2.9	—
MSM/Other							0.2	—
IDU/HC							4.2	8.5
IDU/Other							0.1	0.1
HC/Other							0.1	0.6
MSM/IDU/HC							1.1	—
MSM/IDU/Other							0.02	—
MSM/HC/Other							0.03	—
IDU/HC/Other							0.07	0.2
MSM/IDU/HC/Other							0.02	—

[a]Adjusted for reporting delay

MSM=male-to-male sex

IDU=injection drug use

HC=heterosexual contact

Other includes all confirmed other modes of transmission (e.g., blood/blood products, transplant, health care worker exposure).

NOTES: May not total 100% due to rounding.

A dash (—) indicates not applicable.

A blank indicates not considered.

tributed for HIV cases, the proportion with no transmission risk was set to 0%. The proportion in all other categories increased as those without risk were distributed according to historical patterns. When PHC was computed from extant data, 545 (1.9%) men were considered PHC, reducing the proportion without transmission risk to 30.6%. Among 14,668 women reported with an initial HIV diagnosis, 6,346 (43.3%) were initially reported without transmission risk. Using the standard method, when risk was redistributed for HIV cases, the proportion with no transmission risk was set to 0%. The proportion in all other categories increased as those without risk were distributed according to historical patterns. When PHC was computed from extant data, 450 (3.1%) women were considered PHC, reducing the proportion without transmission risk to 40.2%.

To examine the effect of potential misclassification of PHC, we examined the proportion of individuals with known MSM and IDU risk that would be classified as PHC if they had denied MSM or IDU. Among 40,243 MSM AIDS cases diagnosed from 1999 through 2001, 2,319 (5.7%) would be misclassified as PHC if MSM were denied. Among 17,380 male and 7,610 female IDU cases, 2,228 (12.8%) and 766 (10.1%) would be misclassified as PHC if IDU had been denied. Percentages were similar among HIV diagnoses that had not progressed to AIDS (data not shown).

Table 5. Percentage distribution of HIV transmission risk as reported, after standard redistribution, with the addition of presumed heterosexual contact (PHC), and using mutually exclusive risk categories, HIV diagnoses not progressing to AIDS[a]—1999–2001, 29 HIV-reporting states, U.S., *n*=44,078

	Risk as reported, using current hierarchy		*Risk redistributed, using current statistical redistribution method*		*PHC, using recommended definition*		*Mutually exclusive categories, using recommended categories*	
	Male	*Female*	*Male*	*Female*	*Male*	*Female*	*Male*	*Female*
MSM	44.7	—	60.5	—	44.7	—	40.5	—
IDU	8.3	11.0	15.3	20.0	8.3	11.0	5.5	5.2
MSM/IDU	3.4	—	5.4	—	3.4	—	2.4	—
HC	10.8	45.1	18.2	78.7	10.8	45.1	10.8	44.7
Other	0.4	0.6	0.6	1.3	0.4	0.6	0.4	0.6
Undetermined	32.4	43.3	0.0[b]	0.0[b]	30.6	40.2	30.6	40.2
PHC					1.9	3.1	1.9	3.1
MSM/HC							4.1	—
MSM/Other							0.1	—
IDU/HC							2.7	5.7
IDU/Other							0.1	0.1
HC/Other							0.1	0.4
MSM/IDU/HC							0.9	—
MSM/IDU/Other							0.02	—
MSM/HC/Other							0.03	—
IDU/HC/Other							0.1	0.1
MSM/IDU/HC/Other							0.01	—

[a]Adjusted for reporting delay

[b]Undetermined set to 0.0 in modeling risk redistribution

MSM=male-to-male sex

IDU=injection drug use

HC=heterosexual contact

Other includes all confirmed other modes of transmission (e.g., blood/blood products, transplant, health care worker exposure).

NOTE: May not total 100% due to rounding.

A dash (—) indicates not applicable.

A blank indicates not considered.

Mutually exclusive risk categories

Removal of the hierarchical structure of transmission risk categories results in the addition of 11 categories (Tables 4 and 5). All original hierarchy categories represent a smaller proportion of cases. Each additional category constitutes less than 9% of cases, ranging from 8.5% for female IDU/HC to 0.01% for male MSM/IDU/HC/other.

DISCUSSION

Using currently available information on risk behaviors in the national HIV/AIDS surveillance database, we were able to classify only 1.4% to 3.1% of cases initially classified as no risk reported into a presumed heterosexual category (defined as sexual contact with opposite-sex partner and denial of all other risk factors). Men were less likely than women to be classified as PHC and had a greater chance of being misclassified as presumed heterosexual compared with women.

Our inability to classify more cases into PHC was related to the large proportion of cases with missing or unknown risk behavior data. Among male AIDS cases, for example, 40% had missing or unknown information on sex with a female, 35% had missing or unknown information on sex with a male, and 48% had missing information on IDU. Among female AIDS cases, 21% had missing or unknown sex with male and 52%

had missing information on IDU. The proportion of cases with missing or unknown data in behavior risk fields varied by sex and reporting site (data not shown). Without acknowledgement of sexual contact with an opposite-sex partner and denial of other risk behaviors, PHC could not be assigned. Assignment of NRR to PHC might be more successful in states or local areas where supplemental surveillance projects are conducted, allowing collection of additional information from patient interview[3] and review of medical records at regular intervals.[2] Using supplemental surveillance data in Michigan, Schmidt et al.[5] were able to document additional risk behaviors on reported HIV cases. However, during our analysis period, only eight of the 29 HIV-reporting states used in this analysis were implementing supplemental surveillance activities similar to those available in the analysis from Michigan.[5]

PHC may be less applicable to men, as it is often difficult to rule out MSM, resulting in an undercount of cases attributed to MSM. A multi-site validation study of risk behaviors in the AIDS reporting system in the mid-1990s indicated that heterosexual risk was overestimated for women by up to 29% and for men by up to 70%.[6] In addition, most studies suggest that female to male transmission of HIV is less efficient than male to female transmission.[7]

Removing the risk behavior hierarchy, we added 11 categories. Each additional category constituted a small proportion of cases, and the proportion of cases without risk remained high. Too many categories can be unwieldy for interpretation, especially if the proportions are low in most groups. Presenting these exhaustive, mutually exclusive risk categories is only as effective as data collection is complete. As is the case with assignment of PHC, incomplete risk behavior information in our data limited our ability to assign cases into finer categories. Completeness of data collection for risk behaviors would need to improve for this presentation to be beneficial. However, with improved data collection, this method of presentation would enable us to avoid assumptions about which transmission mode is more likely if more than one is present, to present data as they are collected, and to identify growing categories that might be overshadowed by combining into a broader group.

Our inability to reduce substantially the proportion of NRR cases in the national database simply by creating additional categories or assigning cases to PHC based on currently available data suggests a measurement or information ascertainment issue. The CDC is piloting several efforts to improve ascertainment of transmission risk. Twelve areas have been funded to evaluate and refine a standard protocol that examines existing records to assess the most efficient and cost-effective sources of risk information. In addition, we plan to identify and evaluate barriers faced by health care providers in collecting and documenting HIV transmission risk. Based on identified barriers, we plan to develop interventions to improve providers' ascertainment and documentation of this critical information. We are collecting supplemental high risk behavior information, including exchanging sex for drugs or money, use of illicit non-injection drugs, and recent history of other sexually transmitted infections on a sample of cases with heterosexual contact that does not fit the current CDC hierarchy definition. We will examine whether these data facilitate the creation of a standard definition for high-risk heterosexual behavior. Finally, six high-morbidity areas have been funded to pilot sampling methods allowing us to make inference about transmission risk from a representative sample to the population of all persons with diagnosed HIV disease in these six areas.

A scientific approach using a standard definition of high-risk heterosexual behavior has been recommended over the presumption that cases having heterosexual partners and no other reported risk factor be designated as presumed heterosexual contact.[8] It is clear that effective surveillance of HIV risk behaviors among individuals with newly diagnosed HIV will require a variety of data collection strategies and statistical approaches. The CDC will use the data from these pilot studies to improve the completeness and accuracy of data on transmission risk behaviors and to guide the development of any changes to the HIV transmission mode hierarchy.

The authors thank Mona Saraiya, MD, MPH, and John E. Gerstle III, MS, for their assistance with data preparation, and Danni Daniels, MS, for critical review of an earlier draft of the manuscript.

REFERENCES

1. Nakashima AK, Fleming PL. HIV/AIDS surveillance in the United States, 1981–2001. J Acquired Immune Defic Syndr 2003;32 Suppl 1:68-85.
2. Farizo K, Buehler J, Chamberland M, Whyte B, Froelicher E, Hopkins S, et al. Spectrum of disease in persons with human immunodeficiency virus infection in the United States. JAMA 1992;267:1798-805.
3. Buehler JW, Diaz T, Hersh BS, Chu SY. The supplement to HIV-AIDS surveillance project: an approach for monitoring HIV risk behaviors. Public Health Rep 1996; 111Suppl 1:133-7.
4. Green TA. Using surveillance data to monitor trends in the AIDS epidemic. Stat Med 1998;17:143-54.
5. Schmidt MA, Mokotoff ED. HIV/AIDS surveillance and

prevention: improving the characterization of HIV transmission. Public Health Rep 2003;118:197-204.
6. Klevens RM, Fleming PL, Neal JJ, Li J. Is there really a heterosexual AIDS epidemic in the United States? Findings from a multisite validation study, 1992–1995. Am J Epidemiol 1999;149:75-84.
7. Padian NS, Shiboski SC, Glass SO, Vittinghoff E. Heterosexual transmission of human immunodeficiency virus (HIV) in northern California: results from a ten-year study. Am J Epidemiol 1997;146:350-7.
8. Fleming PL, Jaffe HW. AIDS among heterosexuals in surveillance reports [letter]. N Engl J Med 2001;344: 611-3.