Guidelines for Designing Rapid Assessment Surveys of HIV Seroprevalence Among Hospitalized Patients

BERNHARD SCHWARTLÄNDER, MD
ROBERT S. JANSSEN, MD
GLEN A. SATTEN, PhD
SARA E. CRITCHLEY, RN
LYLE R. PETERSEN, MD, MPH
TIMOTHY J. DONDERO, MD

The authors were all associated with the Public Health Service's Centers for Disease Control and Prevention when these guidelines were developed. Dr. Schwartländer is currently with the Global Program on AIDS, World Health Organization, Geneva, Switzerland. Dr. Janssen, Dr. Satten, Dr. Petersen, and Dr. Dondero are with the Division of HIV/AIDS. Ms. Critchley is with the Hospital Infections Program.

Tearsheet requests to Robert S. Janssen, MD, CDC, Division of HIV/AIDS, MS E46, 1600 Clifton Road, Atlanta, GA 30333; tel. 404-639-2082; FAX 404-639-2029.

Synopsis

The Centers for Disease Control and Prevention has developed guidelines for determining HIV seroprevalence among patients seeking medical care at acute-care hospitals. The guidelines enable hospital staff members to perform a simple, rapid, and inexpensive survey to determine seroprevalence among the patient population, protecting the anonymity of those who are tested. The guidelines are based on national experience with large-scale anonymous, unlinked HIV serosurveys.

The data from a rapid assessment survey are particularly useful for evaluating the need to provide routine, voluntary HIV counseling and testing and treatment for HIV infection. Beyond that, such data can be used in targeting education efforts, in reinforcing the use of appropriate universal precautions, in resource allocation, and in determining the need for further studies of HIV infection among the population in the hospital catchment area.

PATIENTS WITH AIDS have been admitted to more than 90 percent of the metropolitan area hospitals in the United States and to about 40 percent of the rural area hospitals (*1*). In addition, a large number of persons with undiagnosed HIV infection have been hospitalized (*2*). The number and proportion of their patients infected with HIV is unknown in most hospitals, however.

Hospital seroprevalence data have many uses —

- evaluating the need to provide routine, voluntary HIV counseling and testing services as recommended by the Centers for Disease Control and Prevention (CDC) (*3*),
- evaluating the need for treatment services for HIV infection and associated conditions,
- targeting education efforts,
- reinforcing the use of appropriate universal precautions by health care workers,
- assisting in making decisions regarding resource allocation, and
- determining the need for further epidemiologic studies about demographic and behavioral characteristics in the hospital's catchment population.

Based on the experience of the 39 U.S. acute-care hospitals participating in the Sentinel Hospital Surveillance System for HIV Infection (*2,4,5*), CDC has developed and field-tested generic guidelines (*6,7*), approved by a CDC institutional review board for protection of human subjects. Under the guidelines, hospital personnel can perform a simple, rapid, and inexpensive survey to determine the HIV seroprevalence among the patient population and protect the anonymity of those who are tested. The guidelines provide for anonymous and unlinked testing of residual blood specimens collected for routine diagnostic purposes. Tested patients cannot be identified. This approach prevents bias from self-deferral, such as when persons at high risk for HIV infection refuse to participate in a voluntary study (*8*). The guidelines are flexible, so they can easily be adapted to the special needs or interests of a particular hospital.

'The guidelines provide help in addressing important issues involved in performing such a survey, including anonymity of patients, scientific validity, and generalizability of results.'

This is not to imply that hospitals must or should conduct such surveys. Our purpose is to assist those hospitals that wish to determine their HIV seroprevalence. The guidelines provide help in addressing important issues involved in performing such a survey, including anonymity of patients, scientific validity, and generalizability of results. The results from a hospital survey can apply to its treated population but not necessarily to the general population of its catchment area.

It is unclear how frequently these surveys should be performed. Staffs at some hospitals may want to conduct a single survey, others might want to repeat the survey annually or less frequently.

Survey Design

For reasons of ethics and simplicity, exhaustive efforts must be made to destroy any possible link between patients and their blood specimens prior to HIV serologic testing to avoid inadvertent identification of patients whose blood is tested.

Data obtained by CDC's Sentinel Hospital Surveillance System for HIV infection (*2,5*) show an age distribution similar to that of AIDS patients overall, with the highest seroprevalence rates in the 25–44-year-old age group. To obtain maximal precision of seroprevalence estimates with a minimal sample size, we suggest sampling patients ages 15–54. Since the seroprevalence in patients younger or older than this age category is likely to be very low, sampling in children and the elderly would add little information and could increase the chance of inadvertent identification of a patient who was tested, particularly in hospitals where the overall seroprevalence is low (less than 1 percent).

These seroprevalence surveys can be conducted and coordinated by many different persons in a hospital but require knowledge of routine procedures in the hospital's chemistry and hematology laboratories. Possible principal investigators include laboratory directors or infectious disease physicians collaborating with laboratory personnel. Alternatively, infection control personnel can be principal investigators and should collaborate with both laboratory and infectious disease personnel in conducting these surveys. Principal investigators should have knowledge of research study design, issues about confidentiality, and data analysis. In addition, personnel will need to be knowledgeable about laboratory practices, specimen handling, data collection, data management, and HIV-antibody testing.

Demographic data. The distribution of AIDS cases and HIV infection varies widely by geographic area, sex, age, and race (*9,10*). In this protocol, for simplicity, "race" is used to include ethnicity (for example, Hispanic) as a mutually exclusive race category, even though members of the ethnic group may be of various races. The rates are generally higher among men compared with women and among blacks and Hispanics compared with whites. They are highest among people ages 25 to 44. Therefore, we suggest that data about age group, sex, and race be included. (Specific year of age in combination with the other variables might compromise anonymity; precise year of age is not very useful for the analysis and should not be collected.)

Some hospitals may have the ability to collect additional data such as information regarding patient's known HIV status or HIV-related disease. This information could be used to help estimate the number of unrecognized or unsuspected HIV infections in the overall patient population. Collected data considered together, must not, however, allow linkage of HIV test results to specific people.

Sample size. For this survey, the recommended sample size is 1,000 patients. With this sample size, if there is only 1 seropositive out of the 1,000 tested, the approximate 95 percent confidence interval for the true seroprevalence can be construed to be between 0.002 percent and 0.56 percent. In a sample with a 1 percent seroprevalence rate (10 seropositive specimens), the equivalent interval would be 0.48 percent to 1.84 percent. If more precise results are needed, larger sample sizes would be required.

Stratification of the sample. To generalize the results of the sample to a defined, hospital-patient subgroup (hospital inpatients or a particular patient service, for example), the sample must be representative of the patient population. Blood specimens collected sequentially in a given period, such as 1–2 days, may not constitute a representative sample, however. To correct for possible biases, the best sample includes stratification by age and sex to

ensure that the sample consists of the typical age and sex distribution of patients ages 15–54 at the hospital. The stratification also simplifies the avoidance of duplicate specimens, data management, and analysis of data by age and sex.

To obtain the distribution by age and sex of an average patient population ages 15–54 in a specific hospital, discharge data from a representative period, obtained from the hospital medical records or billing department, can be separated by age and sex of all patients in that age category into eight strata: ages 15–24, 25–34, 35–44, and 45–54 for each sex. In general, the discharge data of the last calendar year should yield representative data of an average patient population. Changes like the opening of a new trauma care unit in the middle of the year, however, may have a major impact on the patient distribution. In such a case, one has to be sure that the population from which the age and sex distribution is taken is representative of the actual patient population (for example, discharge data of the last 3 months.)

The proportion of each group is calculated by dividing the number of discharges in that group by the total number of discharges of patients ages 15–54. An example of this procedure for an average hospital is given in table 1. To obtain the number of specimens per age- and sex-stratum that should be included in the survey sample, it is necessary to multiply the total sample size (1,000) by the proportion of the appropriate group (table 2). Although this procedure is not a randomized sampling procedure, it should provide a good approximation of the patient population because the numbers per stratum are calculated in advance. If different populations, such as outpatients and inpatients, are included in the survey, collection of separate samples of 1,000 specimens for each population is recommended.

Sampling problems. Laboratory or hospital procedures can also lead to a sampling bias. For example, if some specimens of a particular patient group such as obstetric patients are sequestered for other purposes like special studies, those specimens may not be available for this survey. When designing the survey and interpreting results, investigators should determine if such procedures are practiced in their laboratories.

In some hospitals, the total sample could be selected in less than 1 week. For example, if a large group of patients is admitted to the hospital on a certain day of the week for a specific procedure like elective surgery, these patients could be either over- or underrepresented in the sample, depending on whether specimens drawn on that particular day were included or excluded from the sample. One way to compensate for such a bias is to select only every second or third available specimen to ensure that sampling occurs over a sufficiently long period.

Table 1. Distribution of an average hospital population by age and sex for a certain time period, such as the last calendar year. (The percentages result from dividing the number of individuals in the age and sex group by the total number of patients ages 15–54 (14,730))

Age group (years)	Men		Women		Total	
	Number	Percent	Number	Percent	Number	Percent
15–24	938	6.4	2,793	18.9	3,731	25.3
25–34	1,208	8.2	3,810	25.9	5,018	34.1
35–44	1,339	9.1	1,844	12.5	3,183	21.6
45–54	1,341	9.1	1,457	9.9	2,798	19.0
Totals	4,826	32.8	9,904	67.2	14,730	100.0

SOURCE: National Hospital Discharge Survey, 1988, National Center for Health Statistics.

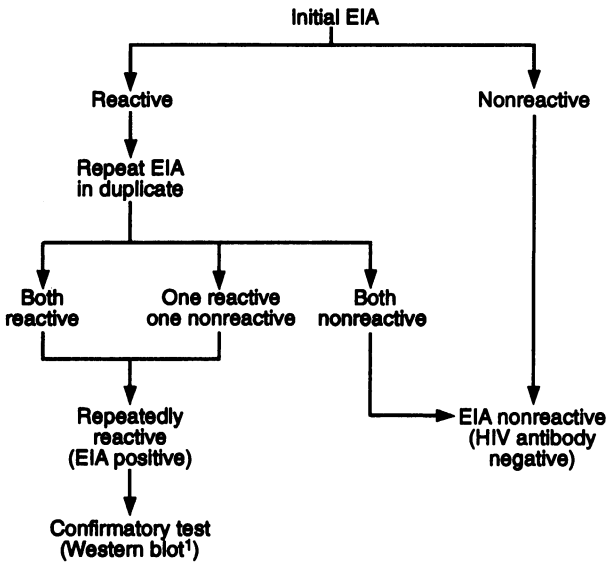
Table 2. Sample sizes for each age and sex group, obtained by multiplying the total sample of 1,000 by the proportion of the corresponding group calculated in table 1

Age group (years)	Men		Women		Total	
	Number	Percent	Number	Percent	Number	Percent
15–24	64	6.4	189	18.9	253	25.3
25–34	82	8.2	259	25.9	341	34.1
35–44	91	9.0	125	12.5	216	21.6
45–54	91	9.1	99	9.9	190	19.0
Totals	328	32.8	672	67.2	1,000	100.0

Sampling Procedure

To perform this survey, investigators need access to residual blood specimens and corresponding demographic data. The most efficient way to perform the survey is to collect blood specimens and the corresponding demographic data at one central laboratory. The chemistry laboratory, for example, might be the most likely place to obtain leftover serum or plasma without additional procedures like centrifugation of heparinized blood and have specimens from patients that meet the survey population requirements. In addition, personnel at the collection laboratory will need access to patient data on age group, sex, race, and clinical service. If hospitals have different laboratories for different units, investigators could either perform separate surveys of the patients of the different units or stratify the

The sequence of serodiagnosis procedures for human immunodeficiency virus using enzyme immunoassay (EIA)



[1]Recommendations for the interpretation and the use of the Western blot assay for serodiagnosis of HIV-1 infections have been published (7).

sample according to the overall patient distribution in the different units. Specimens could then be collected consecutively for each age group and sex stratum in each laboratory.

Data collection form. The form used to collect sample data is designed to enable personnel to collect the data, conduct the stratified sampling while avoiding duplicate testing of the same person, and protect the anonymity of patients being selected. Sampling with this form does not require the use of a computer; however, a computerized data base can be easily developed from it.

Before the beginning of the survey, investigators must clearly define the information they want to collect. The recommended variables are sex, age group, race, and clinical service. Specific categories are given for sex and age group. The race variable has two generally useful fixed categories of white and black patients. Another race-ethnicity (Hispanic, for example) may be specified if the proportion of this group is at least 10 percent of the patient population. No particular race or ethnic group should be used if it represents less than 10 percent of the patient population; all such groups should be combined and classified as "other." For the service unit, there are four proposed categories, surgery, internal medicine, obstetrics, and psychiatry. Other units, if essential to note, should be categorized in advance. In addition, it could be indicated whether a patient is treated as an inpatient, outpatient, or in the emergency room.

If additional information is of special importance to a particular hospital, it may be collected and specified. It is especially important to specify such categories in advance to be sure that the meaning is unequivocal and that the data collection is systematic. Any additional data should be essential to interpreting and using the results of the survey and should not allow for the identification of a person tested in the survey. Additional information, such as specific diagnosis, might allow for inadvertent identification, if the diagnosis is rare. Therefore, investigators must be cautious in collecting any additional information.

Avoiding duplicate sampling. Besides the survey number, which may be preprinted on the data collection forms, the last 4 digits of the patient's hospital identification number (usually 7 digits) can be recorded temporarily on a data collection form to avoid duplicate sampling. These segments of the patient numbers must be removed and destroyed before the specimens are serologically tested. Each specimen should be labelled with the corresponding survey number and prepared labels attached to the data collection forms.

Collection of specimens and corresponding data. When the patients and the laboratory site where specimens will be collected are identified and the total number of patients in each of the eight age and sex strata is determined, data collection forms should be prepared. For example, if all specimens can be collected in one laboratory site and the total number of specimens for men ages 25–34 is 82, four data collection forms (for 25 specimens each) could be labeled for this age-sex stratum and stapled together. The last number for this stratum should be highlighted to make it easy to recognize when the total number of this particular stratum is achieved.

The specimen selectors in the laboratory can then take the following steps:

1. Check each specimen after routine testing to determine whether there is enough serum (or plasma) left over. A volume of 0.5 milliliter (ml) or more would be ideal and would allow for extra tests for indeterminate specimens. A minimum of 0.2 ml is enough for the routine HIV serology, however. If there is less than 0.2 ml available, the specimen should be excluded.
2. If there is at least 0.2 ml available, check whether the specimen fits into one of the eight

sex-age strata (ages 15–54 years). If not, exclude it and check the next specimen. If yes, check whether there is already a specimen included in the corresponding age-sex stratum with the same patient number (last four digits). If yes, exclude it and check the next specimen.

3. If not already sampled (as indicated by the last 4 digits of patient number), the specimen should be selected and the last four digits of the patient number and demographic data recorded on the form. Such data may need to be located through the laboratory's computer system or at another place in the hospital with access to patient demographics. A sample of 0.5 ml (at least 0.2 ml, see step 1) should be taken and the corresponding label, which is attached to the data collection form, should be placed on the tube. No other labels should be attached to the tube holding the sample. Persons aliquoting specimens should wear gloves and a laboratory coat. Some laboratory workers may also prefer to work in a hood, behind a plastic shield, or wear safety glasses (*11*).

4. When all forms for a particular stratum are filled (and, thus, the total number of specimens in this stratum is achieved), continue to collect specimens only in the incomplete strata.

5. When the forms of all eight strata are full and before submitting the specimens for HIV testing, cut off and destroy the columns with the last four digits of the patient number.

Specimen Storage

After the aliquots are collected and the tubes are labeled with their survey numbers, the specimens should be stored in a freezer (at least -20° C) until the total sample size is achieved and the serums (or plasma specimens) are tested for HIV antibodies. If the specimens cannot be frozen the day they are collected, they may be temporarily stored in a refrigerator (4°– 8° C) for a maximum of 5 days.

Guidelines for HIV Antibody Testing

The specimens should be tested (see chart) using an FDA-licensed HIV-1 enzyme immunoassay (EIA) kit. If the specimen is nonreactive in the initial test, the specimen is classified as negative. If the specimen is reactive in the initial test, the EIA should be repeated in duplicate, using fresh aliquots of the same specimen. If at least one of the repeated EIAs is reactive, the specimen is classified as "EIA repeatedly reactive" and should be tested with a supplemental test. The recommended supplemental test is an FDA-licensed Western blot kit. Guidelines for the interpretation and use of the Western blot assay have been published by the Association of State and Territorial Public Health Laboratory Directors and Centers for Disease Control and Prevention (*12*).

'The time required to conduct a rapid assessment of HIV seroprevalence survey depends on familiarity with laboratory-based studies, the size of the hospital, and the number of laboratory specimens available.'

HIV antibody results can be recorded on the data collection form or entered into a computer file. The laboratory technician who performed the HIV antibody testing, but who did not select the specimens to be tested, may enter the HIV antibody results into the appropriate data base. The person(s) who selected the specimens should not see individual test results or the completed data collection forms, as a further protection of the anonymity of the survey.

Data Management

A sequence of preprinted survey numbers on data collection forms and preassigned sticker labels facilitate recording of the laboratory results directly on data collection forms. If desired, data can be transferred to a computer (with a data management package such as EPI-Info). The survey can be conducted without the use of a computer, however.

If a computer is used to avoid duplicate sampling of the same patient, a separate temporary file, containing only the age-sex strata and the last 4 digits of the patients' numbers (and not the survey numbers), should be used during sampling to ensure that no link between patient number and HIV test result is possible. After the sampling is completed and before the HIV tests are performed, the data file with the last 4 digits of the patient number must be deleted.

If a computer is used for data analysis, a permanent file with survey number (but without the last 4 digits of the patient number) and demographic information should be used for entering HIV antibody test results. The permanent data file should not contain any information that links

Table 3. 95 percent confidence limits for a Poisson parameter for surveys with less than 21 HIV-positive specimens[1] (regardless of the size of the sample)[2]

Observed number of HIV-positive specimens	Lower bound	Upper bound
0	0.00	3.00
1	0.02	5.57
2	0.24	7.23
3	0.62	8.77
4	1.09	10.24
5	1.62	11.67
6	2.20	13.06
7	2.81	14.42
8	3.45	15.73
9	4.12	17.09
10	4.80	18.39
11	5.49	19.68
12	6.20	20.96
13	6.92	22.23
14	7.65	23.49
15	8.40	24.74
16	9.15	25.98
17	9.90	27.22
18	10.69	28.45
19	11.44	29.67
20	13.00	30.89

[1] The limits of the 95 percent confidence interval for a prevalence (in percent) are calculated by the value of the respective bounds times 100 percent divided by the sample size. For example, if 13 seropositive specimens are found in a sample of 91 (prevalence = 14.3 percent), the 95 percent confidence interval is 6.92 X 100 percent ÷ 91 = 7.60 percent (lower bound); 22.23 X 100 percent ÷ 91 = 24.43 percent (upper bound).
[2] Reference *13*.

Table 4. Number of specimens per stratum and number found to be HIV seropositive, using hypothetical data of the type that will be obtained in a survey using this protocol and examples of the interpretation of results[1]

Age group (years)	Men		Women		Total	
	Number	Positive	Number	Positive	Number	Positive
15–24	64	2	189	2	253	4
25–34	82	11	259	11	341	22
35–44	91	13	125	5	216	18
45–54	91	6	99	2	190	8
Totals	328	32	672	20	1,000	52

[1] To obtain the calculation of seroprevalence estimates for a single stratum of men ages 35–44, divide 13 by 91 and multiply the result by 100 percent. The answer is 14.3 percent. In samples with less than 21 positive specimens, a 95 percent confidence interval is 7.60, 24.43, using the formula in table 3. The overall seroprevalence of this population would be 5.2 percent (52 X 100 percent, divided by 1,000). An approximate 95 percent confidence interval for the values of the seroprevalence estimate using the formula on page 000 for samples with more than 20 positive specimens would be 3.97, 6.82.

personal identifiers, such as patient numbers, with HIV test results.

The principal investigator can use the data base for analysis of aggregated data but should not report line-by-line results. The principal investigator must be responsible for ensuring that inadvertent linkage of HIV antibody test results to a specific patient does not occur. For further protection of the anonymity of the patients, the person(s) collecting the specimens and checking patient records should not see line-by-line HIV test result data. The specimen collector(s) may, however, see grouped summary data.

Interpretation of the Results

The estimate of seroprevalence in the patient population is

$$P = \frac{x}{n}$$

where x is the total number of seropositive patients observed (summed over all strata) and n is the total sample size (number tested). (Multiplying by 100 will express this seroprevalence in percent positive.) Estimates of seroprevalence in a subgroup may also be obtained by dividing the number of seropositives from that subgroup by the number tested in that subgroup. For example, separate estimates of seroprevalence for men and women may be calculated.

Approximate confidence limits for the seroprevalence can be obtained with the Poisson distribution. Thus, confidence limits for seroprevalence in the individual strata or groups of strata may be calculated by using the procedures described in table 3 if 20 or fewer HIV-seropositive specimens are observed.

If more than 20 HIV-seropositive specimens are observed, regardless of the size of the sample, the 95 percent confidence interval is calculated by using the following equation:

$$\frac{x + 1.92 \pm 1.96\sqrt{x + 0.96}}{n} \text{ X } 100 \text{ percent}$$

where x = observed number of positive specimens and n = sample size.

As an example, suppose that 21 people are found to be seropositive out of a sample of 275 (seroprevalence = 7.6 percent). Then the 95 percent confidence interval for the seroprevalence is

$$\frac{21 + 1.92 \pm 1.96\sqrt{21 + 0.96}}{275} \text{ X } 100 \text{ percent}$$

In this case, the 95 percent confidence interval is 4.99 percent, 11.67 percent.

An example of the interpretation of results

obtained by a survey following this protocol is given in table 4.

Personnel Time

The time required to conduct a rapid assessment of HIV seroprevalence survey depends on familiarity with laboratory-based studies, the size of the hospital, and the number of laboratory specimens available. In addition to the principal investigator's time, the equivalent of one person working full time for 1 to 2 months is needed to complete the specimen collection and handling and data management. In addition, time is needed for a technician to perform HIV-antibody testing.

References

1. Lewis, C. E., and Montgomery, K.: The HIV-testing policies of U.S. hospitals. JAMA 264: 2764–2767, Dec. 5, 1990.
2. Janssen R. S., et al.: HIV infection among patients in U.S. acute-care hospitals: strategies for the counseling and testing of hospital patients. N Engl J Med 327: 445–452, Aug. 13, 1992.
3. Recommendations for HIV testing services for inpatients and outpatients in acute-care hospital settings. MMWR Morb Mortal Wkly Rep 42 (No. RR-2): 1–6, Jan. 15, 1993.
4. St. Louis, M. E., et al.: Methods of surveillance for HIV infection at U.S. sentinel hospitals. Public Health Rep 105: 140–146, March-April 1990.
5. St. Louis, M. E., et al.: Seroprevalence rates of human immunodeficiency virus infection at sentinel hospitals in the United States. N Engl J Med 323: 213–218, July 26, 1990.
6. Murrill, C. S., et al.: HIV-1 seroprevalence in hospital patients in rural Georgia. South Med J 185: 969–971 (1992).
7. Rhodes, L. V., Reed, J. F., Faust, L. A., and Moser, K.: Rapid assessment of human immunodeficiency virus seroprevalence in a community-based hospital. Am J Infect Control 21: 9–15 (1993).
8. Pilot study of a household survey to determine HIV seroprevalence. MMWR Morb Mortal Wkly Rep 40: 1–5, Jan. 11, 1991.
9. Centers for Disease Control: HIV/AIDS surveillance. Year-end edition. February 1993.
10. National HIV serosurveillance summary: results through 1990. Division of HIV/AIDS, National Center for Infectious Diseases, Centers for Disease Control, November 1991.
11. Centers for Disease Control: Guidelines for prevention of transmission of human immunodeficiency virus and hepatitis B virus to health-care and public-safety workers. DHHS Publication No. (NIOSH) 89–107, U.S. Government Printing Office, Washington, DC, 1989.
12. Interpretation and use of the Western blot assay for serodiagnosis of human immunodeficiency virus type 1 infections. MMWR Morb Mortal Wkly Rep 38: 1–7, July 21, 1989.
13. Odeh, R. E., Owen, D. B., Birnbaum, Z. W., and Fisher, L.: Pocket book of statistical tables. Dekker, New York, 1977.