AIDS Rev 2002;4:213-223

Street and Network Sampling in Evaluation Studies of HIV Risk-Reduction Interventions

Salaam Semaan[1], Jennifer Lauby[2] and Jon Liebman[3]

[1]National Center for HIV, STD, and TB Prevention, Centers for Disease Control and Prevention, Atlanta, Georgia, USA. [2]Philadelphia Health Management Corporation, Philadelphia, Pennsylvania, USA. [3]Holyoke Health Center, Holyoke, Massachusetts, USA

Abstract

Although sampling is a crucial component of research methodology, it has received little attention in intervention research with populations at risk for HIV infection. We review the challenges involved in sampling these populations for evaluating behavioral and social interventions. We assess the four strategies used for street and network sampling that have been reported in the HIV-intervention research literature and used because traditional probability sampling was not possible. The sampling strategies are: 1) targeted, 2) stratified, (3) time-space, and (4) respondent-driven. Although each has strengths and limitations in terms of its ability to produce valid results that enhance generalizability, the choice of a particular strategy depends on the goal of the study, characteristics of the target population, and the availability of resources and time for collecting and analyzing sampling-related data. Continued efforts are needed to improve the sampling strategies used in evaluation studies of HIV risk-reduction interventions.

Key words

Survey sampling. HIV. AIDS. HIV prevention. Rare populations.

Introduction

The HIV/AIDS epidemic is a pressing public heath issue, and valid scientific information is needed to guide community prevention efforts. The purpose of HIV-prevention research is to learn which interventions, when widely implemented, will reduce infection and disease[1]. Thus, the generalizability of study results to different populations and settings is critical. However, the generalizability of results is dependent on high-quality sampling, because sample selection and composition are important considerations for the replication and dissemination of intervention studies.

Although probability sampling is broadly the ideal for facilitating the generalizability of results, selecting probability samples of populations at risk for HIV infection in order to evaluate behavioral and social interventions is often difficult and expensive. Street and network sampling strategies have been used in community-based intervention studies to minimize biases associated with non-probability sampling and to select either probability samples or representative samples of populations at risk for HIV infection.

Sampling options

Samples can be categorized into three broad categories on the basis of how they were selected. In probability sampling, each person of the

Correspondence to:
Salaam Semaan
National Center for HIV, STD, and TB Prevention
Centers for Disease Control and Prevention
Mailstop E-02, 1600 Clifton Rd., Atlanta, GA 30333
Email: SSemaan@cdc.gov

target population has a known, non-zero probability of selection, and such sampling allows inferences to the population on the basis of standard statistical theory[2]. Probability sampling is regarded as ideal because standard statistical methods permit researchers to make inferences from the sample to the target population. Probability sampling provides a means of decreasing investigator- and respondent-associated biases, evaluating the reliability of estimates and the magnitude of the sampling error, and producing unbiased estimates that can be generalized to the target population[2-4].

In representative samples, researchers cover a broad cross-section of the target population to encompass its heterogeneity, and important characteristics are distributed similarly in both the sample and the population[5]. Representative samples are ideal when strict probability samples cannot be selected, and they are preferable to non-probability samples.

Non-probability sampling provides valuable information for problem definition[6,7], but poses difficulties in generalizing study results. Non-probability samples may be biased by unknown magnitudes and direction as a result of how the samples were selected. Results from biased samples improperly influence conclusions about prevalence estimates of important variables and influence inferences regarding relationships between variables[8]. Examples of non-probability sampling are quota sampling, in which a certain number of participants from certain demographic or behavioral subgroups are interviewed, and convenience sampling, in which accessible participants are interviewed. Although convenience sampling has been used in many HIV-intervention studies, several calls were made in the early 1990s for innovative procedures that would yield samples that were representative of the target populations[9-11].

The probability-sampling model of generalization is the most satisfactory model for explaining when and how results can be generalized[2]. However, important as this model may be to enhancing external validity and the generalizability of results, selecting probability or representative samples of people at risk for HIV infection may be prohibitively costly in terms of time and money. In response to these challenges, an alternative model has been used in generalizing from study results obtained in community-based HIV interventions to the target population that is the focus of an intervention[12]. This alternative model is based on the heterogeneity, or the diversity principle[13,14]. According to this model, researchers are able to generalize a study effect from similar interventions (e.g., small-group interventions) implemented with heterogeneous populations (e.g., male drug users, women at risk for HIV infection) in different settings (e.g., a sexually transmitted disease [STD] clinic, a prison). Thus, the diversity model necessitates a reasonable number of studies of similar interventions, conducted with diverse populations in diverse settings. Researchers in HIV intervention have relied mostly on the diversity model rather than on the sampling model in identifying effective interventions for transfer to programs[12]. Reasons for the dependence on the diversity model include the need to work through the challenges that are specific to community-based studies[15] and those that are specific to the target populations, and the topics of investigation and the resources available for sampling-related activities.

Community-based HIV intervention studies

In response to the HIV epidemic, numerous evaluation studies of behavioral and social interventions have been implemented[16,17]. These risk-reduction intervention studies provide a promising and crucial approach to the prevention of HIV infection[17,18]. More specifically, the interventions aim to change individual-level sexual or drug-related risk behaviors and to change peer-related or community-wide norms and practices to support individuals' efforts to adopt risk-reduction behaviors[19,20]. Typically, researchers conduct the interventions with populations at high risk for HIV infection in geographically defined areas. In these interventions, researchers employ street or network sampling methods, apply behavioral and social science theories, and use experimental or quasi-experimental designs with concurrent comparison groups[20]. Researchers use HIV risk behaviors and infection rates as outcome measures to evaluate the effect of these interventions [for example[21-28]].

Internal and external validity

Important methodological dimensions of intervention studies, whether they are randomized clinical trials or quasi-experimental designs with concurrent comparison groups, are the internal and external validity[29].

Internal validity refers to the absence of substantial differences between study groups at baseline. It is ensured either by using random assignment or an unbiased, systematic method of assignment to study groups. It is also crucial that there be no substantial differences in attrition between the study groups at follow-up. Many evaluations of HIV prevention interventions have used randomization, matching, or an unbiased, systematic method to select participants[17]. These assignment methods ensure that biases in sampling are the same in both study groups (ensuring high internal validity) and allow confidence in measuring the intervention effect. They do not, however, ensure external validity because they do not require a representative or a probability sample of the population.

External validity, or the ability to make inferences from the study sample to the target population, is a much more difficult issue and requires that the sample be a probability sample or a representative sample[2]. In recent years, medical research has attempted to increase the diversity of study samples in clinical trials, to improve the ability to generalize from study results to the target population[30,31]. In clinical trials of medical interventions, there is now more emphasis on including women and men of various ages, racial backgrounds, and stages of risk behaviors and disease[32,33]. These developments highlight the importance of external validity and the need for samples that are representative of the target population and for results that are generalizable to the target population.

In the evaluation of behavioral on social interventions such as HIV prevention interventions, the representativeness of the sample is important because the acceptance of the risk-reduction intervention messages depends on the norms, attitudes, and experience of the participants and the target population. These determinants of behavior change are dependent on factors such as gender roles, racial and ethnic identification, and sociodemographic variables. A key part of the evaluation is to adequately sample the community and assess the proportion of the population reached by the intervention. A randomized clinical trial of an intervention that works well in a self-selected sample may not be effective if it is disseminated to the target population. Therefore, representative or probability samples are particularly important in evaluating community-based interventions that aim to provide risk-reduction messages to everyone at risk in the targeted communities.

Challenges specific to community-based studies

In some studies, *community* has been defined as a geographic area of residents and non-residents who frequent the area to transact business or to seek entertainment. In other studies, *community* has been defined as membership in a certain population; for example, a population characterized by sexual or drug-related risk behaviors[34,35]. Operationally, these definitions of a community do not facilitate constructing a sampling frame (a comprehensive list) of the target population or determining its size and characteristics. Not surprisingly, it is equally difficult to develop a sampling frame of the non-residents frequenting a geographically defined area, to estimate their number, and describe their characteristics[36]. The unknown mobility of members of the target population does not help to ensure that each member has a known, non-zero probability of selection. Thus, constraints in defining the community make it difficult to delineate geographic boundaries, to determine the denominator of the target population, and to assess the probability of selecting the study sample.

Changing community conditions also add to the complexity of constructing and updating sampling frames and reproducing the strategy at other sites and at other times, especially in multiyear intervention studies that are evaluated by analyzing data collected from repeated cross-sectional samples[37,38]. For example, houses may be demolished, businesses may close, and commercial sex workers or drug users may change their territory in response to police pressure, community action, or other conditions[39]. Other practical constraints also affect the ease of selecting representative or probability samples. These constraints include weather conditions, safety concerns, periodicity of attendance of the study population at accessible locations, and negotiating with community leaders for access to the community. These conditions also influence interview rates, including people's availability for street interviews and the difficulty of rescheduling anonymous street interviews, and they have implications for external validity.

Challenges specific to the population and topics of investigation

Unlike the four strategies described in this article, traditional probability sampling (e.g., door-to-door household surveys, telephone surveys, and facility-based surveys) underestimates the population at risk for HIV infection because it is less likely to reach these populations (e.g., drug users and commercial sex workers)[40,41]. In many instances, the cost of screening the general population to construct a sampling frame of high-risk persons, or to estimate the size of the high-risk population, may exceed the cost of interviewing these high-risk persons[42]. In addition, the time needed to construct a sampling frame may be impractical, even unethical, in view of the urgent need to control the HIV epidemic[43]. Some members of these populations may not live in conventional dwellings or seek health care regularly, and many may provide unreliable answers to protect their privacy if selected through traditional probability strategies[44,45]. The use of public data sets does not facilitate the selection of people at high risk for HIV infection because these data sets often do not capture the population of interest (e.g., commercial sex workers).

Populations at high risk for HIV infection may be considered rare, because they constitute a small proportion of the general population[40], hidden, because they are difficult to find[36,40], and difficult to sample, because their size is difficult to determine[46]. Kish[47] defined *elusive populations* as those for which it is difficult to construct a sampling frame that allows the application of probability sampling.

The topics of investigation in HIV prevention research, which necessitate questions about high-risk sexual and drug-related behaviors, add to the challenges of selecting representative or probability samples. These topic-related challenges stem from the private nature of sexual behaviors, the illegality of many behaviors, particularly drug use, and the social stigma associated with engaging in HIV risk behaviors and with becoming HIV infected[40,48].

Sampling in HIV intervention studies

To address the various challenges of sampling encountered in community-based HIV intervention studies, researchers have used several alternative strategies for selecting street and network samples of the population at risk for HIV infection. Some of these strategies are better than others for eliminating, reducing, or controlling bias. In some of these strategies, random selection is used in several, but not all, stages of sampling, and the exact probability of selecting each person is unknown. Although less than ideal from a theoretical perspective, these strategies are a practical necessity for selecting representative samples when traditional sampling methods are not likely to yield successful results.

Information on sampling strategies appears in reports of study outcomes and in reports of methods. We reviewed the methodological reports published during 1985-2001 of the sampling strategies used in community-based HIV-prevention intervention research. We obtained the reports by searching databases (AIDSLINE and Medline) and networking with other researchers. As part of another project, researchers are reviewing the sampling information in outcome reports (CDC, *unpublished data*, 2000). We found four distinct strategies used in place of traditional probability sampling for selecting street and network samples in community-based HIV intervention studies: 1) targeted, 2) stratified, 3) time-space, and 4) respondent-driven.

Targeted sampling

Targeted sampling calls for the collection and use of quantitative data (e.g., institutional data) and qualitative data (e.g., ethnographic interviewing) in selecting the sample[49,50]. Existing *indicator data*, usually collected by public or private agencies for surveillance or service-delivery needs, are used to delineate the geographic boundaries for the intervention and to describe the target population. Indicator data typically include admissions to drug treatment, drug-related hospital emergency admissions, drug-related deaths, drug-related arrests, cases of STDs and HIV/AIDS, and population characteristics from census data. Additional data sources include ethnographic data and the observations of outreach workers. These data are used to develop the sampling frame of locations where the target population may be found and to characterize the sample in terms of variables such as age, race, location, and other important features. Specific locations for sampling are then randomly selected from the sampling frame. Interviewers then go to the selected locations and recruit drug users who are willing to participate in the study or interview.

Targeted sampling has been used in community-based HIV interventions to select street samples of out-of-treatment drug users[36,39,41,51-53]. Targeted sampling depends on the accuracy and comprehensiveness of ethnographic mapping and the analysis of indicator data. It is clearly better than convenience or other forms of nonprobability sampling because it ensures the inclusion of persons with varied demographic and risk characteristics. Targeted sampling requires experienced staff and several months of data collection and analysis before recruiting and interviewing can begin.

Targeted sampling has several limitations. First, because indicator data were collected for other aims, they may not be very useful in characterizing the target population, or even the population that needs HIV-prevention interventions. For example, poor and minority populations tend to be over-represented in police statistics, or data may be reported for different geographic areas such as police districts (drug-related arrests), health districts (STDs), or zip codes (AIDS cases and drug-related hospital emergency admissions).

Second, because targeted sampling does not specify how persons are selected for recruitment, the resulting sample may be biased and difficult to replicate. Interviewer and self-selection biases may result when random selection rules are not used to sample drug users. For example, because of safety concerns, outreach workers may restrict the recruitment of drug users to highly visible public areas or may place themselves where drug users know where to find them[55,56]. As a result, targeted sampling may undersample drug users who do not occupy obvious niches in the community or who do not approach outreach workers (e.g., young injection drug users and those who maintain daily activities, such as jobs, despite their drug use)[57]. The analysis of longitudinal data has shown that targeted sampling tends to recruit those who are at higher risk for HIV infection, and only later to recruit those at lower risk[58].

Finally, in targeted sampling, geographic areas may not be sampled in proportion to numbers of drug users, drug users may not be sampled in proportion to intensity of drug use, and the probability of selecting drug users may not be known. All of these factors influence the validity or generalizability of the data[2,54].

Stratified sampling

In stratified sampling, the target community is divided into strata; for example, community locations (e.g., houses, residential blocks, street corners, hangouts, parks), businesses (e.g., bars and restaurants), and agencies (e.g., shelters and drop-in centers). To prepare for the sampling process, project staff members list the locations in each stratum and then construct a sampling frame of the target population for only the locations that they have randomly selected from the list. Thus, the sample is selected from each stratum in stages. In stage one of a two-stage sampling strategy, a random sample of the locations is selected from each stratum. In stage two, researchers select a sample of the target population from the randomly selected locations. Kipke, et al.[37,59] used this strategy to select a sample of high-risk street youth. This strategy was also used at the Philadelphia site of the Women and Infants Demonstration Project to evaluate an HIV risk-reduction intervention with young, sexually active women[11,60,61].

In stratified sampling, researchers conduct formative research and collect sampling-related data to define the target population and characterize the strata, and then use these data to determine the proportions of the sample selected from each stratum. This strategy allows researchers to construct a sampling frame of the target population for the last stage only (i.e., for the randomly selected locations), thus eliminating the need for, and the impracticality of, constructing sampling frames for all strata and reducing sampling-related costs. If a detailed sampling frame of people cannot be constructed for the randomly selected locations, predefined, strata-specific random or systematic selection rules can be used to reduce interviewer- and self-selection biases. Stratified sampling allows researchers to concentrate the project staff members in the selected locations rather than to spread the field staff throughout the community, also reducing sampling-related costs. Stratification allows researchers to analyze data by relevant characteristics of the population (e.g., locations of high-risk vs low-risk populations) and to produce stratum-specific estimates. In addition to the selection of local community residents, stratified sampling also allows the selection of non-residents who frequent the selected locations and people who engage in high-risk activities – two groups that are common targets of community-based HIV interventions. An added advantage is that researchers can replicate the sampling procedures to study changes in behavior. An important caution is the issue of intraclass correlation, which arises when the persons at a location have homogeneous characteristics. Researchers need to incorporate this effect (also known as the clustering or design effect) into the statistical analysis to obtain valid inferences[2,62-66].

Difficulties with stratified sampling include keeping the list of the locations up-to-date, especially as community conditions change, covering all strata and subgroups of the population, sampling accurate proportions of high-risk groups, particularly when their size is not known, and accounting for participants' mobility between locations. Because it may be difficult to weight the sample from each stratum, the sample may not be representative of the target population.

Time-space sampling

Time-space sampling (also known as venue-based sampling) is a probability-based method for recruiting members of a target population at specific times at congregating locations (also known as venues)[38,67]. Venue-day-time (VDT) units, (e.g., a 4-hour period on Monday at a given location), representing the universe of potential places, days and times, form the sampling frame. Project staff members identify the range of VDT units for the members of the target population by interviewing service providers, key informants, and members of the target population. Following that, the project staff members canvass the venues and prepare a list of VDT units that are considered potentially eligible on the basis of headcounts and screening interviews. This information allows project staff members to estimate for each VDT unit the size of the population and the number eligible for sampling. The sample is then selected in stages. In stage one of a two-stage sampling strategy, project staff members select a simple random or stratified sample of all VDT units listed on the sampling frame (preferably with a probability proportional to the total number of the eligible population for each VDT unit). In stage two, participants are systematically selected from the randomly selected VDT units. Time-space sampling also allows sampling in informal venues, such as house parties or parades, to reach closeted or younger members of the target population, or those who typically do not frequent the more public venues. Time-space sampling was used to select a representative sample of young men who have sex with men (MSM) in the Community Intervention Trial for Youth[68,69] and was also used in survey research with MSM[38,70,71].

Time-space sampling depends on the formative research on which it is based and is more practical for lengthy, well-funded studies. An accurate count of the persons at the venues is extremely important, and researchers need to understand the attendance habits of their populations. Time-space sampling provides for the selection of a probability sample or a representative sample of the target population who visit the venues listed in the sampling frame. To allow generalization to the target population, visitors and non-visitors of the listed VDT units, researchers need to use a weighting mechanism to reflect the probability of attendance for the VDT units listed on the sampling frame, the probability of selecting a specific VDT unit, and the probability of being interviewed for the selected VDT unit[38]. To ensure a probability sample, project staff members need to sample all entrances of spacious venues in proportion to frequency of use.

Time-space sampling assumes that most of the target population has set patterns for visiting different locations, which may be true for some populations, but not for others. Excluding venues or times of low attendance or excluding venues that provide limited access may introduce selection biases. Oversampling smaller venues may ensure a more representative sample, but this procedure may be expensive[69].

Respondent-driven sampling

Respondent-driven sampling is predicated on the recognition that peers are better able than outreach workers and researchers to locate and recruit other members of a hidden population. At First glance, respondent-driven sampling may seem similar to snowball sampling, because both snowball and respondent-driven sampling are types of chain-referral sampling. However, they do differ in major ways.

Snowball sampling is a form of non-probability sampling in which participants give the study staff members the names of their peers so that staff members can recruit them into the study. Because snowball sampling is not based on probability theory, researchers cannot produce population estimates based on the data collected through such sampling. Respondent-driven sampling is designed to produce probability samples of the target population and to overcome the biases associated with snowball sampling[57,72,73]. Like standard probability sampling, respondent-driven sampling provides a means of selecting a sample and evaluating the reliability of the data obtained, and allows inferences about the characteristics of the population from which the sample is drawn[72].

Similar to other forms of chain-referral sampling, respondent-driven sampling starts with a small number of peers, usually chosen non-randomly. However, unlike the recruitment procedures used in other forms of chain-referral sampling, the procedures in respondent-driven sampling incorporate the direct recruitment of peers by their peers, recruitment quotas (e.g., three recruits only), and a dual system of incentives (for participation and for recruiting). Respondent-driven sampling was used in evaluating an HIV risk-reduction intervention with drug users[74]. "Steering" (i.e., additional) incentives have been used to increase the recruitment of peers that belong to specific subgroups, such as injection drug users who are 18 to 25 years old[73]. Respondent-driven sampling is particularly useful for sampling populations that do not trust the research community[75]. Analysis of data obtained from respondent-driven samples of drug users has shown that respondent-driven sampling controls the biases associated with chain-referral sampling and results in representative samples of drug users[57,72,73].

The strengths of respondent-driven sampling are many. Quotas lengthen referral chains, helping to eliminate or reduce the biases associated with the initial choice of peers, and allowing deeper penetration into the diverse and isolated sectors of the population[57]. Quotas also reduce volunteerism (the willingness to participate as a respondent) and limit the ability of peers with larger personal networks to recruit more extensively than peers with smaller networks[57]. The choice of the recruitment quota depends, in part, on the size of the population and the sample size required. Recruitment quotas ensure that semiprofessional recruiters do not merge and battle for turf, thus giving all peers the opportunity to recruit their peers. The incentive structure and the quota system reduce masking (the lack of willingness to refer peers for participation as respondents) because these procedures allow participants to ask their peers whether they are willing to participate, and allow participants who may not be comfortable giving the name of a peer to a researcher to recruit that person directly. Refusal by potential participants is a problem inherent in all sampling strategies because participation in research studies is always voluntary. However, respondent-driven sampling allows researchers to determine empirically the extent to which volunteerism affects sampling by examining the association between self-reported network composition and peer-recruitment patterns.

To ensure the recruitment of members who differ from the initial recruiting peers, and to ensure an adequate sample of the population, at least four to six chains (waves) of recruitment are needed. This principle of recruitment, validated empirically for respondent-driven sampling, is based on the "six degrees of separation", which suggests that only half a dozen links are required to connect everyone in the country[76]. Procedures have been established for computing the number of waves needed to obtain a representative sample of the target population[57,73]. To assess the adequacy and biases of the sample, researchers can compare the data collected from participants on the composition of their personal networks (e.g., demographic and behavioral characteristics) and the characteristics of those who were actually recruited. The rather close correspondence between the characteristics of the persons recruited and the characteristics of their networks (recruitment matrix) has shown the potential of respondent-driven sampling for reducing volunteerism and masking[73]). Additionally, analytic procedures have been established for weighting the sample to compensate quantitatively for differences in homophily the tendency to recruit peers with characteristics similar to those of the recruiters and network size[73].

Respondent-driven sampling assumes that persons at risk are connected in a network and have ties to other persons at risk. Because respondent-driven sampling does not require large-scale retrieval and analysis of indicator data, it proceeds relatively quickly. Formative research is embedded in the sampling process and interview data and is not required before sampling starts. Respondent-driven sampling provides the ability to select a sample and to collect information about the social structure from which the sample is drawn. As community conditions (e.g., network structure) change, researchers can analyze the recruitment matrix data to identify these changes.

Respondent-driven sampling provides the ability to reach persons who shun public venues and is practical in small, modestly-funded projects. It is particularly appropriate when sampling groups of people who are linked by certain behaviors, such as drug use, or certain characteristics, such as HIV infection.

Respondent-driven sampling has several limitations. Although it is easier to collect and compare data on the socio-demographic characteristics and the network composition of persons recruited and those of the recruiters, it is harder to do so for the more personal HIV-related characteristics, such as sexual and drug-related behaviors and HIV status. These characteristics are a concern because they may be associated with infection probabilities[77-79]. Recent work in respondent-driven sampling has shown its potential to produce representative samples, although these results need to be replicated to gain wider acceptance by the research community.

Conclusions

Studies evaluating community-based HIV interventions need to be conducted with scientific rigor to allow confidence in the results and generalizability of results to the target population. However, generalizability of the results is often dependent on the hard-to-implement traditional probability sampling strategies, which cannot usually be implemented successfully in community research settings, particularly when studying high-risk sexual and drug-related behaviors. Thus, although traditional probability sampling is important to reduce selection biases and to enhance the validity and generalizability of results, it is difficult and expensive to conduct in community-based HIV intervention studies. These challenges make the use of creative sampling strategies to select representative or probability samples of high-risk populations a real necessity. Although we focused on the challenges faced and the sampling strategies used in community-based HIV interventions, these challenges and strategies are applicable to sampling other hard-to-reach populations in community health studies[42,46,80-85]. We focus our concluding remarks on a comparison of the sampling strategies reviewed, the importance of training staff in key sampling procedures, and ways to advance sampling strategies in community-based HIV intervention studies.

The sampling strategies we reviewed have the potential to produce representative or probability samples when sampling-related data are collected and analyzed. These strategies also offer pragmatic options for maintaining a high degree of external validity in evaluation studies of community-based HIV risk-reduction interventions without being overly cumbersome or costly. Although some of these strategies are stronger than others for selecting representative or probability samples, they differ in important logistical and analytical features. We present a summary of these features (Table 1) that can be used as a guide in the selection of sampling strategies for intervention studies.

Targeted, stratified, and time-space sampling all use trained staff and a listing process to select the sample, but other aspects of the strategies differ. The strategies differ in the amount of formative research and preparation needed, in the duration and budget required for sampling activities, in the extent of the sampling-related biases, and the methods by which those biases are controlled. Targeted sampling builds on community characteristics and available data, uses formative research, and requires fewer resources than do the other strategies. Stratified sampling allows the inclusion of community residents and non-residents, incorporates a wide range of community locations, and is less costly than time-space sampling. Time-space sampling is good for selecting regular visitors to specific community locations and for sampling groups that congregate in accessible venues. It allows selection of a representative sample and offers the opportunity to select a probability sample. Time-space sampling is less suitable if a significant part of the target population does not congregate in public venues. In such instances, respondent-driven sampling is more suitable. Respondent-driven sampling uses the least amount of formative research and resources, uses peers to recruit the sample, and has been used to sample a hidden subgroup of drug users. Whereas in targeted sampling ethnographers and staff members must spend considerable time gaining the trust of participants, in respondent-driven sampling respondents are recruited by acquaintances, so sampling proceeds more quickly.

Researchers need to use appropriate analytic tools that take into consideration sampling-related data (e.g., control for selection biases, intra-class correlation, and weighting). In stratified sampling and in time-space sampling, for example, the use of weights controls for differential selection probabilities[2]. In intervention studies that use peers to recruit the sample and to deliver the intervention, it is important to differentiate between the intervention effect associated with the peer-delivered intervention and the effect associated with peer recruitment. Thus, researchers may need to examine whether the peer-sampling strategy is itself a low-level intervention that enhances the effect of the peer-delivered intervention.

Because sampling procedures are especially tedious in community settings, staff training and intensive quality-control measures are needed. In table 2, we present selected sampling steps that could be used in sampling strategies that use trained staff to select the sample. These steps capitalize on the strengths of the sampling strategies we reviewed and can be used as criteria for evaluating the sampling procedures used in community-based intervention studies.

Although our review reveals that researchers have used particular sampling strategies with cer-

Table 1. *Summary of selected features of sampling strategies*

Strategy	Population		Formative research	Study duration	Budget	Recruiters	Strengths or Limitations	External Validity
	Reported	Recommended						
Targeted	Drug users	Street	Considerable	Multiyear	Limited	Staff	Excludes certain locations, times, and populations	Limited
Stratified	Youth at risk Women at risk	In multiple settings (e.g., community locations, businesses, and agencies)	Considerable	Multiyear	Moderate	Staff	Excludes certain times (e.g. evening hours) Includes community residents and nonresidents	Moderate
Time-space	Men who have sex with men	In public locations	Considerable	Multiyear	Considerable	Staff	Selects visitors to specific locations Provides a representative or probability sample	Considerable
Respondent -driven	Drug users Young drug users	Rare populations[a] In private locations Inaccessible, reclusive	Minimal	Short	Minimal	Peers	Provides a representative or probability sample	Considerable

Note: The sampling strategies have the potential to produce representative or probability samples when sampling-related data (e.g., weights) are collected and analyzed.
[a]Rare populations constitute a small proportion of the general population.

tain populations (e.g., targeted and respondent-driven sampling with drug users, time-space with MSM, and stratified with heterosexual populations), these sampling strategies are not specific to certain populations. The choice of a particular sampling strategy depends on study aims, characteristics of the target populations, and considerations of external validity. In selecting a sampling strategy, researchers need to consider the advice of sampling statisticians, the biases (interviewer, self-selection, location, and time) that are inherent in the strategy, and the various ways that these biases can be avoided or minimized through formative research and through collection and analysis of sampling-related data. To improve sampling strategies in HIV intervention research, it is important that researchers and editors of scientific journals and books value the reporting of details of the sampling strategies used. It is equally important that authoritative books on sampling cover the sampling of hidden populations.

The theoretical and methodological benefits of representative or probability samples are clear, but the effects of different sampling strategies on research results are harder to demonstrate. Several questions warrant further consideration. First, we know little about the comparative costs of the strategies or the comparative representativeness of samples of the same population. Sampling research may benefit from comparing the cost and composition of samples of the same population (e.g., drug users or MSM) selected from a given community by different sampling strategies. It would be important, for example, to know whether some members of the population would be missed in time-space sampling but selected in respondent-driven sampling. Alternatively, in some studies researchers may use more than one strategy, e.g., targeted and respondent-driven sampling, to select the sample and obtain the required sample size. It seems important that we also develop an understanding of whether combining strategies in a single study would bring out complementary strengths or complementary weaknesses of the different strategies used, or would make it hard to interpret the sampling design that was used and the results that were obtained.

Second, researchers can answer sampling-related questions by analyzing relevant data sets and using modeling or meta-analytic techniques. Because internal validity does not ensure external validity, it is crucial to examine whether intervention studies that use internally valid methods produce different effects depending on the sampling strategies used. Researchers can analyze intervention effects in relation to sampling strategies and in relation to population characteristics. Such analyses may be useful in learning about the situations in which sampling biases are most likely to affect the overall conclusions of a study. In other words, researchers can examine whether there is a relationship between significant intervention effects and sizes and the general rigor of the sampling strategy. Thus, it may be possible to

Table 2. Sampling steps
Before sampling
1. Include in the sampling frame all strata and locations (sites or venue-day-time units) 2. Approximate the number of the target population in each stratum and location 3. Determine proportional allocation of sample between different strata and locations 4. Train interviewers to use the sampling strategy 5. Implement ways to boost participation rates in the screening and core interviews
During sampling
1. Select a random or a proportionate stratified sample of locations from each stratum 2. Choose each location with a probability proportional to the estimated total of the target population 3. Use random or systematic sampling when feasible, or use random selection procedures to select respondents 4. Interview all persons who are eligible 5. Collect data on mobility of respondents (mobility affects the probability of selection) 6. Collect data to calculate weights in order to control for an unequal probability of selection 7. Collect data to assess selection biases in the screening and core interviews 8. Supervise the sampling
During analysis
1. Compute attendance rates at locations and rates of participation in the screening and core interviews 2. Compare respondents interviewed in the different strata and locations, by characteristics associated with the outcomes of interest 3. Assess reasons for refusal to participate in the screening and core interviews and determine whether refusal rates are associated with selection biases 4. Assess representativeness of the selected samples by comparing the data with other data sets (e.g., a similar strategy in a different setting or a different strategy in a similar setting) 5. Incorporate weights into the analyses to reflect unequal probabilities of selection, incomplete sampling frame, and rates of no response 6. Assess the need to use statistical programs that incorporate the design effect 7. Compare the achieved sample and the desired sample

examine whether a trend in intervention effects is associated with the sampling strategy used.

Finally, future studies can examine the effects of using representative or probability samples by comparing intervention effects for different segments of the sampled population, or by weighting the sample to favor easy-to-sample groups and comparing results when hard-to-reach populations are included. Modeling exercises [for example[86]] can be used to compare a particular sampling strategy and a more traditional sampling strategy with respect to bias and variability of estimates. Triangulation, or the use of several data sets, can be used to evaluate the extent to which a sample is representative of the target population.

For the foreseeable future, stopping the spread of the HIV epidemic among high-risk populations relies in part on the development, evaluation, and transfer of effective HIV risk-reduction interventions[87,88]. It is crucial to use sampling strategies that control various sampling biases and to continue to improve selection of representative or probability samples of the full spectrum of the target population. These efforts would increase the scientific rigor, credibility, and usefulness of community-based HIV intervention studies. Although populations at risk for HIV infection may be referred to as "difficult to sample", probability or representative samples from these populations can be selected with good planning, adequate resources, and the collection and analysis of sampling-related data.

Acknowledgments

We thank John E. Anderson, James W. Carey, Melissa Cribbin, Doug Heckathorn, Lilian Lin, Greg Millet, and Ann O'Leary for their insightful comments on an earlier version of the manuscript.

References

1. Neumann M, Sogolow E, Holtgrave D. Supporting the transfer of HIV prevention behavioral research to public health practice. AIDS Educ Prev 2000;12(Suppl A):1-3.
2. Levy P, Lemeshow S. Sampling of Populations: Methods and Applications. New York: John Wiley and Sons 1991.
3. Bobbie R. Survey Research Methods. Belmont, CA: Wadsworth Publishing 1973.
4. Jaeger R. Sampling in Education and the Social Sciences. New York: Longman 1984.
5. Spreen M, Zwaagstra R. Personal network sampling, outdegree analysis and multilevel analysis: Introducing the network concept in studies of hidden population. International Sociology 1994;9:475-91.
6. Bernard H. Research Methods in Anthropology: Qualitative and Quantitative Approaches. Thousand Oaks, CA: Sage 1994.
7. Patton M. Qualitative Evaluation and Research Methods. Newbury Park, CA: Sage Publications 1990:187-9.
8. Leigh B, Stall R. Substance use and risky sexual behavior for exposure to HIV: Issues in methodology, interpretation and prevention. Am Psychol 1993;48:1035-45.
9. Coyle S, Boruch R, Turner C (eds). Methodological issues in AIDS surveys [appendix C]. In: Evaluating AIDS Prevention Programs (expanded ed.). Washington, DC: National Academy Press 1991:207-316.

10. Coyle S, Boruch R, Turner C (eds). Sampling and randomization: Technical questions about evaluating CDC's three major AIDS prevention programs [appendix D]. In: Evaluating AIDS Prevention Programs (expanded ed.). Washington, DC: National Academy Press 1991:317-34.
11. Miller H, Turner C, Moses L (eds). AIDS: The Second Decade. Report for the NRC Committee on AIDS Research and the Behavioral, Social, and Statistical Sciences. Washington, DC: National Academy Press 1990.
12. Hedges L, Johnson W, Semaan S, Sogolow E. Theoretical issues in the synthesis of HIV prevention research. J Acquir Immune Defic Syndr 2002;30(Suppl 1):8-14.
13. Cook T. A quasi-sampling theory of the generalization of causal relationships. New Directions in Program Evaluation 1993;57:39-82.
14. Campbell D, Fiske D. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull 1959;56:81-105.
15. Susser M. The tribulations of trials – intervention in communities. Am J Public Health 1995;85:156-60.
16. Auerbach J, Coates T. HIV prevention research: Accomplishments and challenges for the third decade of AIDS. Am J Public Health 2000;90:1029-32.
17. Sogolow E, Peersman G, Semaan S, et al. The HIV/AIDS prevention research synthesis project: Scope, methods, and study classification results. J Acquir Immune Defic Syndr 2002;30(Suppl 1):15-29.
18. Des Jarlais D, Semaan S. HIV prevention research: Cumulative knowledge or accumulating studies? An introduction to the HIV/AIDS Prevention Research Synthesis project supplement. J Acquir Immune Defic Syndr 2002;30(Suppl 1):1-7.
19. Kelly J. Community-level interventions are needed to prevent new HIV infections. Am J Public Health 1999;89:299-301.
20. Semaan S, Kay L, Strouse D, et al. A profile of U.S.-based trials of behavioral and social interventions for HIV risk reduction. J Acquir Immune Defic Syndr 2002;30(Suppl 1):30-50.
21. Kelly J, Murphy D, Sikkema K, et al. Randomized, controlled, community-level HIV prevention intervention for sexual risk behavior among homosexual men in U.S. cities. Lancet 1997;350:1500-5.
22. Johnson W, Hedges L, Ramírez G, et al. HIV prevention research for men who have sex with men: A systematic review and meta-analysis. J Acquir Immune Defic Syndr 2002;30(Suppl 1):118-29.
23. Lauby J, Smith P, Stark M, Person B, Adams J. A community-level HIV prevention intervention for inner-city women: Results of the Women and Infants Demonstration Projects. Am J Public Health 2000;90:216-22.
24. Mullen P, Ramírez G, Strouse D, Hedges L, Sogolow E. Meta-analyses of the effects of behavioral HIV prevention interventions on the sexual risk behavior of sexually experienced adolescents in U.S. trials. J Acquir Immune Defic Syndr 2002;30(Suppl 1):94-105.
25. Neumann M, Johnson W, Semaan S, et al. Review and meta-analysis of HIV prevention intervention research for heterosexual adult populations in the United States. J Acquir Immune Defic Syndr 2002;30(Suppl 1):106-17.
26. Needle R, Coyle S, Normand J, Lambert E, Cesari H. HIV prevention with drug-using populations – current status and future prospects: Introduction and overview. Public Health Rep 1998;113(Suppl 1):4-18.
27. CDC AIDS Community Demonstration Projects Research Group. Community-level HIV intervention in 5 cities: Final outcome data from the CDC AIDS community demonstration projects. Am J Public Health 1999;89:336-45.
28. Semaan S, Des Jarlais D, Sogolow E, et al. A meta-analysis of the effect of HIV prevention interventions on the sex behaviors of drug users in the United States. J Acquir Immune Defic Syndr 2002;30(Suppl 1):73-93.
29. Cook T, Campbell D. Quasi-Experimentation: Design and Analysis Issues for Field Settings. Chicago: Rand McNally 1979.
22. Bernard HR. Research Methods in Anthropology: Qualitative and Quantitative Approaches. Thousand Oaks, CA: Sage 1994.
30. National Institutes of Health. NIH guidelines on the inclusion of women and minorities as subjects in clinical research. Federal Register 1994;59:14508-13.
31. Food and Drug Administration (FDA) modernization act of 1997 (FDAMA or the ACT) 1997:105-15.
32. Gifford A, Cunningham W, Heslin K, et al. Participation in research and access to experimental treatments by HIV-infected patients. N Engl J Med 2002;346:1373-82.
33. King T. Racial disparities in clinical trials. N Engl J Med 2002;346:1400-2.
34. Kelly J, Murphy D, Sikkema K, Kalichman S. Psychological interventions to prevent HIV infection are urgently needed: New priorities for behavioral research in the second decade of AIDS. Am Psychol 1993;48:1023-34.
35. Koepsell T, Wagner E, Cheadle A, et al. Selected methodological issues in evaluating community-based health promotion and disease prevention programs. Annual Review of Public Health 1992;13:31-57.
36. Watters J, Biernacki P. Targeted sampling options for the study of hidden populations. Social Problems 1989;36:416-30.
37. Kipke M, O'Connor S, Palmer R, Mackenzie. Street youth in Los Angeles: Profile of a group at high risk for HIV infection. Arch Pediatr Adolesc Med 1995;149:513-9.
38. MacKellar D, Valleroy L, Karon G, Lemp G, Janssen R. The Young Men's Survey: Methods for estimating HIV seroprevalence and risk factors among young men who have sex with men. Public Health Rep 1996;111(Suppl 1):138-44.
39. Liebman J, Thomas C. Development of a sampling strategy to monitor trends in HIV risk behavior in out-of-treatment drug users. In: Epidemiologic Trends in Drug Abuse: Proceedings of the Community Epidemiology Work Group, June 1992. Rockville MD. Alcohol, Drug Abuse and Mental Health Administration 1992:494-501. DHHS Publication (ADM) 92-1958.
40. Kalton G. Sampling considerations in research on HIV risk and illness. In: Ostrow DG, Kessler RC (eds). Methodological Issues in AIDS Behavioral Research. New York: Plenum Press 1993:53-74.
41. Watters J. The significance of sampling and understanding hidden populations. Drugs and Society 1993;7:13-21.
42. Sudman S, Sirken M, Cowan C. Sampling rare and elusive populations. Science 1988;240:991-6.
43. Laumann E, Michael R, Gagnon J, Michaels S. The Social Organization of Sexuality: Sexual Practices in the United States. Chicago: University of Chicago Press 1994.
44. Catania J, Gibson D, Chitwood D, Coates T. Methodological problems in AIDS behavioral research: Influences of measurement error and participation studies in studies of sexual behavior. Psychol Bull 1990;108:339-62.
45. Gagnon J, Lindenbaum S, Martin J, et al. Sexual behavior and AIDS. In: Turner CF, Miller H, Moses L (eds). AIDS: Sexual Behavior and Intravenous Drug Use. Washington, DC: National Academy Press 1989:150-3.
46. Lepkowski J. Sampling the difficult-to-sample. J Nutr 1991;121:416-23.
47. Kish L. Taxonomy of elusive populations. Journal of Official Statistics 1991;7:340-7.
48. Martin J, Dean L. Developing a community sample of gay men for an epidemiologic study of AIDS. American Behavioral Scientist 1990;33:546-61.
49. Bluthenthal R, Watters J. Multimethod research from targeted sampling to HIV risk environments. In: Lambert EY, Ashery RS, Needle RH (eds). Qualitative Methods in Drug Abuse and HIV Research. Rockville MD. National Institute of Drug Abuse 1995:212-30. DHHS Publication 95-4025. Research Monograph 157.
50. Wiebel W. Sampling issues for natural history studies including intravenous drug abusers. In: Hartsock P, Genser SG (eds). Longitudinal Studies of HIV Infection in Intravenous Drug Users: Methodological Issues in Natural History Research. Rockville MD. Alcohol, Drug Abuse, and Mental Health Administration 1991:51-62. DHHS Publication (ADM) 91-1786. Research Monograph 109.
51. Braunstein M. Sampling a hidden population of non-institutionalized drug users. AIDS Educ Prev 1993;5:131-9.

52. Carlson R, Wang J, Siegal H, Falck R, Guo J. An ethnographic approach to targeted sampling: Problems and solutions in AIDS prevention research among injection drug and crack-cocaine users. Human Organization 1994;53:279-86.
53. Robles R, Colon H, Freeman D. Copping areas as sampling and recruitment sites for out-of-treatment crack and injection drug users. Drugs and Society 1993;7:91-105.
54. Cochran W. Sampling Techniques. New York: John Wiley and Sons 1997.
55. Broadhead R, Fox K. Occupational health risks of harm reduction work: Combating AIDS among injection drug users. In: Albrecht GL, Zimmerman R (eds). Advances in Medical Sociology: The Social and Behavioral Aspects of AIDS. Vol 3. Greenwich, CT: JAI Press 1993:23-42.
56. Johnson J, Williams M, Kotarba J. Proactive and reactive strategies for delivering community-based HIV prevention services: An ethnographic analysis. AIDS Educ Prev 1990; 2:191-200.
57. Heckathorn D. Respondent-driven sampling: A new approach to the study of hidden populations. Social Problems 1997;44:174-99.
58. Iguchi M, Bux D, Lidz V, Kushner H, French J, Platt J. Interpreting HIV seroprevalence data from a street-based outreach program. J Acquir Immune Defic Syndr 1994;7:491-9.
59. Kipke M, O'Connor S, Nelson B, Anderson J. A probability sampling for assessing the effectiveness of outreach for street youth. In: Greenberg JB, Neumann MS (eds). What We Have Learned from the AIDS Evaluation of Street Outreach Projects: A Summary Document. Atlanta, GA: Centers for Disease Control and Prevention 1998:17-28.
60. Semaan S, Lauby J, Liebman J, Walls C, Lavelle K. Opportunities and challenges of community-level sampling strategies. Paper presented at the annual conference of the Eastern Evaluation Research Society. Conshohocken, PA 1996.
61. Semaan S, Liebman J, Walls C. Sampling design and challenges for HIV prevention community-based surveys. Paper presented at the Pennsylvania Public Health Association Annual Scientific Meeting. Harrisburg, PA 1994.
62. Brogan D. Pitfalls of using statistical software packages for sample survey data. In: Armitage P, Colton T (eds). Encyclopedia of Biostatistics. New York: John Wiley and Sons 1998.
63. Donner A, Donald A. Analysis of data arising from a stratified design with cluster as unit of analysis. Stat Med 1987;6:43-52.
64. Koepsell T, Martin D, Diehr P, et al. Data analysis and sample size issues in evaluations of community-based health promotion and disease prevention programs: A mixed-model analysis of variance approach. J Clin Epidemiol 1991;44:701-13.
65. Murray D. Design and analysis of community trials: Lessons from the Minnesota Heart Health Program. Am J Epidemiol 1995;142:569-99.
66. Rooney B, Murray D. A meta-analysis of smoking-prevention programs after adjustment for errors in the unit of analysis. Health Educ Q 1996;23:48-64.
67. Lemp G, Hirozawa A, Givertz D, et al. Seroprevalence of HIV and risk behaviors among young homosexual and bisexual men: The San Francisco/Berkeley Young Men's Survey. JAMA 1994;272:449-54.
68. Muhib F, Linn L, Stueve A, et al. A venue-based method for sampling hard-to-reach populations. Public Health Rep 2001;116(Suppl 1):216-22.
69. Stueve A, O'Donnell L, Duran R, Doval A, Blome J. Methodological issues in time-space sampling in minority communities: Results with Latino young men who have sex with men. Am J Public Health 2001;91:922-6.
70. Stall R, Hoff C, Coates T, et al. Decisions to get HIV tested and to accept antiretroviral therapies among gay/bisexual men: Implications for secondary prevention efforts. J Acquir Immune Defic Syndr Hum Retrovirol 1996;11:151-60.
71. Valleroy L, MacKellar D, Karon J, et al. HIV prevalence and associated risks in young men who have sex with men. JAMA 2000;284:198-204.
72. Heckathorn D. Respondent-driven sampling II: Deriving valid population estimates from chain-referral samples. Social Problems 2002;49:11-34.
73. Heckathorn D, Semaan S, Hughes J, Broadhead R. Extensions of respondent-driven sampling: A new approach to the study of injection drug users aged 18-25. AIDS and Behavior 2002;6:55-68.
74. Broadhead R, Heckathorn D, Weakliem D, et al. Harnessing peer networks as an instrument for AIDS prevention: Results from a peer-driven intervention. Public Health Rep 1998; 113(Suppl 1):42-57.
75. Broadhead R. Hustlers in drug-related AIDS prevention: Ethnographers, outreach workers, injection drug users. Addiction Research and Theory 2001;9:545-56.
76. Watts D. Small Worlds: The Dynamics of Networks Between Order and Randomness. Princeton, NJ: Princeton University Press 1999.
77. Klovdahl A, Potterat J, Woodhouse D, Muth J, Muth S, Darrow W. Social networks and infectious disease: The Colorado Springs Study. Soc Sci Med 1994;38:79-88.
78. Potterat J, Woodhouse D, Rothenberg R, et al. AIDS in Colorado Springs: Is there an epidemic? AIDS 1993;7:1517-21.
79. Rothenberg R, Potterat J, Woodhouse D, et al. Choosing a centrality measure: Epidemiologic correlates in the Colorado Springs Study of Social Networks. Social Networks 1995; 7:273-97.
80. Fantuzzo J. Prevalence and effects of child exposure to domestic violence. Domestic Violence and Children 1999;9:21-31.
81. Faugier J, Sargent M. Sampling hard to reach populations. J Adv Nurs 1997;26:790-7.
82. Ghani A, Donnelly C, Garnett G. Sampling biases and missing data in explorations of sexual partner networks for the spread of sexually transmitted diseases. Stat Med 1998;17:2079-97.
83. Kanouse D, Berry S, Duan N, et al. Drawing a probability sample of female street prostitutes in Los Angeles County. J Sex Research 1999;36:45-51.
84. Kleinman L, Freeman H, Perlman J, Gelberg L. Homing in on the homeless: Assessing the physical health of homeless adults in Los Angeles county using an original method to obtain physical examination data in a survey. Health Serv Res 1996;31:533-49.
85. Mohr W. Family violence: Toward more precise and comprehensive knowing. Issues in Mental Health Nursing 1999;20:305-17.
86. Lemeshow S, Tserkovnyi A, Tulloch J, et al. A computer simulation of the EPI survey strategy. Int J Epidemiol 1985;14:473-81.
87. O'Leary A, DiClemente R, Aral S. Reflections on the design and reporting of STD/HIV behavioral intervention research. AIDS Educ Prev 1997;9(Suppl A):1-14.
88. Sumartojo E, Carey J, Doll L, Gayle H. Targeted and general population interventions for HIV prevention: Towards a comprehensive approach. AIDS 1997;11:1201-9.