6

Sampling Issues in Drug Epidemiology

Colin Taylor and Paul Griffiths

COLIN TAYLOR AND PAUL GRIFFITHS • European Monitoring Centre on Drugs and Drug Addiction (EMCDDA)

1. INTRODUCTION

"*A sample is a group of subjects selected from a larger group in the hope that studying this smaller group (the sample) will reveal important things about the larger group (the population).*" This definition, taken from the 'Dictionary of statistics and methodology: a non-technical guide for the social sciences.' (Vogt 1993) highlights the central issue for sampling: that researchers study a small, known group but in general want to find out about an unknown, larger population. To support the hope that such inferences can be made, drug epidemiologists are faced with the same statistical and technical issues in the selection of the smaller group from the larger as researchers in other areas of social investigation, but unlike many other areas they are faced with a number of additional specific challenges that arise from the very nature of illicit drug use.

At one end of its spectrum illicit drug use is a covert and stigmatized behavior, and the other end of the spectrum, reaches into social acceptability and even common practice. Injecting drug users and opiate or crack-cocaine users are often termed a 'hidden population' or a 'hard-to-access' population, i.e., not only can they be relatively rare in general population terms and unwilling to participate in, or even hostile to, research activities, but as an identifiable group they are not readily accessed through administrative records. The concern of this chapter is to describe the issues of sampling this harder to reach end of the spectrum, working with the sub-group of known drug users who are at any one time in contact with, or come to the attention of, some sort of official body. At the 'softer' end, cannabis users by contrast are a less marginalized, relatively prevalent and more accessible group; in socially accepting settings, they may be more easily studied using more traditional survey techniques, especially when using methods that have been developed to help reduce both response and non-response biases.

The chapter considers how standard sample selection methods have been adapted and developed to overcome the difficulties that hard drug use presents. Sampling methods are usually designed to help with the quantification of behaviors; i.e., they aim to allow descriptions by counts and percentages to be meaningful, and the concerns here are principally with inferences in quantitative research methods. But the circumstances in which quantification is possible are often limiting and much of the research moves beyond aiming to quantify responses or characteristics in a population to more qualitative research approaches that do not

usually attempt to make any statistically based inferences from their observations to the larger population. We note though that these same sampling procedures will often also benefit qualitative research that looks to generalize beyond its immediate subjects.

2. MOVING FROM STANDARD TO MORE INNOVATIVE SAMPLE SELECTION PROCEDURES

Prevalence estimation is one major area of interest for drug abuse researchers, but many studies need to access samples of drug users to explore numerous other topics such as the natural history and career of drug using, help seeking behavior, and the relationship of drug use to offending and to the development of health or social problems. In doing so, innovative adaptations to traditional random sampling procedures are required to overcome the difficulties inherent in selecting samples from this population group.

To understand how these techniques differ, three traditionally standard steps are underscored that enable a researcher to study scientifically the epidemiology, etiology and other characteristics of a disease, or of any behavior in the target population from which the observed sample group is drawn. These steps are crucial to traditional sampling methods because they can improve the chances that the sample selected for observation is representative of the target population and, while they cannot guarantee representativeness, they will allow the researcher to quantify how likely the sample is to be representative on any specified characteristic. Taken together, these steps form strong controls on the traditional selection procedures:

1. Define the target population and construct a list that includes all its members—sometimes this list can be constructed dynamically (as in multi-stage sampling, described below).
2. Select members from the complete list in a controlled, probabilistic way—while equal probability selections may be made, techniques may be necessary that stratify or boost differentially the selection, particularly if there are not many target individuals on the list, but the probability of selection from the list must remain known.
3. Find and interview the sample members, eliciting responses carefully using appropriate tools, measures and procedures—if not possible to interview them discreetly and confidentially in person, alternative ways of collecting the information can be used.

3. DEFINITIONS ARE DIFFICULT

The question of how to define the target population is probably more complex in the drug addiction field than in many others, as discussed more fully in Chapter 1.

It needs to be emphasized that the question of being a 'current drug user' is not the same sort of question as being 'HIV positive' for example, or as having some specific medical diagnosis. While these sorts of questions are not unique to drug abuse epidemiology, answering them does constitute a major task in a well designed study (e.g., Buster et al., 2001). While diagnostic criteria that apply to drug abuse and dependence have been developed under both the International Classification of Diseases 10 and the Diagnostic and Statistical Manual IV, these are often difficult to apply in many research settings in order to define target populations (e.g., Cottler et al., 2001). Usually studies use definitions based on simple behavioral measures of use of some specified drug type over time, with life-time prevalence (ever-use), last month prevalence, and days used in the last month being the most commonly used constructs.

Typically, studies focus on sub-groups of the overall general population, and target definition becomes crucial. For example, for the European Union a target group definition of drug use based on lifetime prevalence of any controlled substance would imply a target between 20 percent and 25 percent of the adult population, whereas a target population defined as current drug injectors would represent probably under 0.5 percent of the adult population of Europe. (EMCDDA, 2003)

Age restrictions are common in drug studies and are relatively easy to apply. As drug-taking behavior is most prevalent in the 15–35 age group, studies may focus exclusively on this range in order to conserve resources. Although drug use may be important outside this age range, studying it across its full age spread would be resource-intensive, making age-group definitions an important design factor.

A common feature of many drug studies is an emphasis on local populations. Drug-taking is often more prevalent and more easily studied in cities, for example, than at a national level. Using a local population presents definitional difficulties that need to be addressed carefully. For instance, determining what is a sensibly defined prevalence rate in a city might dramatically change as a result of broadening the definition of 'city limits'. Where people live is a question that is relatively easy to document if they have a permanent address but many drug users do not. Whether to include those who work or are receiving drug abuse treatment in the designated area, even if they live elsewhere, are issues that need to be given consideration and definitional precision, as these decisions can have a considerable effect on the overall size and the characteristics of the chosen target population.

4. GENERAL POPULATION SAMPLING FRAMES ARE OFTEN NOT USEFUL

Traditional survey methods for estimating levels of any behavior in a population, i.e., population or household surveys, are the standard ways for measuring prevalence. Indeed they can be very effective in monitoring the use of common and

legal substances such as tobacco and alcohol, and to some extent cannabis. They are ineffective, however, at measuring the prevalence of rare, more covert, more stigmatized and more problematic forms of drug use such as injecting heroin or crack-cocaine use. There are numerous ways in which sampling bias can occur in this arena. For instance, injecting drug users or crack-cocaine users are less likely than non-problematic drug users to live in households that are included in general household surveys, injectors and crack-cocaine users may be less likely to agree to be surveyed if asked, and injecting may be less likely to be reported than other forms of drug use practices.

In addition to potential bias, further complications arise from the low prevalence of problem drug use within general population sampling frames, since any sample drawn will often identify so few users that reports of current injecting behaviors may be all but non-existent. Whether higher prevalence rates mean that other drug use, such as the use of cocaine (powder) or ecstasy, could be the subject of general surveys is a moot point (e.g. Degenhardt et al., in press). Survey data in both the United States and Europe show changes in use over time similar to information derived from indirect indicators of drug use (such as seizures and treatment demand). However the extent to which survey data directly reflects underlying trends for these substances remains unclear and other factors, such as changing social attitudes to drug use, also may be important. Survey techniques, such as the use of booster samples, have been developed to increase the numbers of drug users recruited into the sample. Chapter 7 reviews how survey techniques themselves have been adapted to deal with these difficulties. Nonetheless, for researchers wanting to study the behavior of those consuming drugs like heroin or crack cocaine or who are injecting drug users, a general population sampling frame as a starting point is not likely to be either a practical or a methodologically sound option.

4.1. The Usefulness of Schools Surveys

One example of using a restricted sampling frame with definitional difficulties that is worthy of special note is schools surveys. Schools surveys are widely used for exploring exposure to drug use at a specific age or grade level and represent a valuable tool for tracking trends in exposure to drug use over time. School surveys though do have the same limitations as population surveys in respect of hard drug use. As a rule, to take account of the setting for the data collection, specially designed multi-stage cluster samples are used. In essence it is difficult to do other than collect a complete ‘classroom-full of data’ rather than to sample on an individual basis (UNODC, 2003). Furthermore, the sampling frame used in a multi-stage procedure can be simplified, in that listings of the classes that could be chosen need only be made for selected schools. For more information on school surveys refer to the European School Project on Alcohol and Drugs (Hibbell et al.,

2000), Monitoring the Future (Johnston et al., 2004), and to the Health Behavior in School-Aged Children (Currie et al., 2004).

4.2. Using Non-Standard Sampling Frames

School surveys, as described above, are one example of using (as in many two-stage sample surveys) dynamic sampling frame lists, i.e., lists that are never constructed in entirety but only in parts, as needed, in this case in a hierarchic fashion. Dynamic list construction is frequently used in drug abuse studies.

A close parallel to dynamic list construction is the use of area (quadrat) samples and line transect samples. The underlying principle of these approaches is using the physical, geographic layout of the population as a 'sampling frame', selecting clusters of people for the sample through their physical location rather than through a list. Such methods differ from more standard ones only in that the researcher does not know what fraction of the population is obtained in the sample, but in preserves the same statistical properties as those of a standard multi-stage or cluster sample. This method is particularly useful in countries or regions of a country where administrative records are poor or inadequate for compiling a complete sampling frame in the usual manner. This method, discussed more fully below, emerges in drug abuse epidemiology as 'site sampling' and is a major contributor to the development of sampling methods (McKeganey et al., 1992).

5. INFERENCES, SAMPLING CONTROL, BIASES AND ERRORS

One of the principal functions of using a comprehensive sampling frame in classical statistical probability sampling is to enable the researcher to have control over who enters and does not enter the sample. This may sound paradoxical, since the point is to obtain a *randomly derived* sample of the target population, but in fact this is achieved by selecting according to careful probability-based procedures. It should be emphatically distinguished from a *haphazard sample* into which entry is left to chance factors with no control. In such a sample there are always openings for large potential biases to occur, unknown to the researcher, under the whims of fate, other people, and other unidentified factors that determine what type of person is sampled. For instance, these factors may vary from a bias towards people who are available on a Tuesday afternoon (or whenever selection is carried out), who may be of a cheerful and pleasant disposition or of a compliant personality, and so forth. With a controlled sample selection, there is no causal connection between an individual's characteristics and whether or not that person is in or out of the sample; selection depends instead on the toss of a coin, for example, or some other random procedure.

The benefits that flow from probability sampling concern the degree to which the final sample is representative of the target population, focusing on measuring potential "unrepresentativeness." Indeed the whole point of statistical inference, estimation, reliability and confidence intervals can be regarded as quantifying how likely and to what degree the sample is representative of the target population, and using a known probabilistic random sampling procedure is central to these probability calculations.

5.1. What Can Be Done with Non-Probabilistic Samples?

There are two types of non-probabilistic samples that have been commonly used for studying drug users because of their ease of implementation and the apparent directness of their sampling approach. The first type comes under a broad heading of 'convenience samples'. These samples are not constructed using probabilistic representation of the target population, but simply from convenience of access. This might mean the selection of a group of users who are at hand when the research is being planned or carried out or it may mean interviewing any of the researcher's contacts who are known to exhibit the target behavior. Whatever the access route, these samples provide almost no possibility of making broader inferences beyond those individuals immediately studied. These types of samples or study groups fall outside the scope of standard statistical analysis and inference because their ascertainment does not allow the usual generalizations from any analysis of them. These groups could possibly be classified under a heading of 'special populations' and useful insights have been gained from studying them in an explorative context.

The second type of non-probabilistic sample that offers more surface plausibility is one that is self-selected, i.e., the researcher's control over who enters the sample is relinquished and instead people who volunteer are used as a sample. A common procedure to recruit such a group is to advertise in a magazine, on a radio program, or through clubs that are known to have a high percentage of drug users among their readership, audience or clientele. Using this method of recruitment presents a number of biases such as special personality traits, drug-taking behavior, or demographic characteristics that will dramatically influence the findings from such a sample (Winstock et al., 2001; Inciardi and Harrison, 2000). It can be argued that even using a properly constructed approach, respondents need to agree to be interviewed and can refuse to take part, and so in some sense are still considered volunteers. It is well accepted, however, that should such biases, as measured by the number of refusals or more generally the non-response rate, become significant the study results can be severely compromised. It is worth noting that in a survey using a probability sample, the non-response rate can usually be calculated and it is often possible to determine what type of individual is failing to respond, so that corrections can be made for any inherent bias in the statistical analyses.

A third type of sampling is often used in epidemiological experimental and quasi-experimental studies, where the researcher selects a study group of drug abusers that is not intended to represent a larger population but rather to contain individuals who highlight some important characteristic. Many drug studies fall under this category; for example, drug injectors may be recruited who attend a low threshold treatment center and their injecting practices may differ when compared with those who attend specialist out-patient centers. For the most part these studies employ some sort of randomization of a factor of interest, such as type of treatment administered and in a technical sense the role of randomization becomes central. It is not possible though to use results from these studies in wider epidemiological inferences if the study groups have not been properly sampled. The resulting study conclusions are not treated as estimates in a population, but only as comparisons within the study group.

Much of what is done in estimating prevalence of problem drug use and in drawing drug user samples for study is an attempt to approximate random sampling procedures in order to acquire some sort of representative sample of drug users beyond those in contact with services or known to the authorities. Site sampling and chain referral sampling techniques are two vital components in these methods. The field of drug abuse epidemiology is a mixture of, on the one hand, well-formed analyses on special sub-populations that are in themselves usually targeted primarily because they can be sampled randomly, and on the other hand, a set of procedures that try to approximate random sampling of broader populations that are of interest.

5.2. Statistical and Non-Statistical Inferences—Moving From Best to Next Best

The benefits of using random sampling procedures are that the researcher is able to generalize beyond the 'known observations' in the sample and make probabilistic inferences to the whole target population from which the sample has been drawn. While highly desirable, this statistical ideal is not always practically achievable for many important topics of interest. Therefore a second type of inference can be made based on common knowledge and common sense non-statistical grounds. Moving outside statistical inference in this way means that certainty and uncertainty and the extent to which reliance can be placed on the conclusions is not quantifiable, only arguable. These approaches are important in all fields of behavioral research, moving beyond epidemiological results into a more interpretive arena, where such inferences are frequently necessary; and the drug research arena is no exception.

Many important studies in the drug abuse field have provided useful evidence even when probabilistic inference to the larger population could not be shown. For example, in most circumstances drawing an adequate sampling frame

of drug injectors may be difficult to achieve. Nonetheless, studies based on injectors recruited in streets and other settings have yielded useful information on issues such as injecting practices that have extremely important implications for public health policy. In the design of such studies, drug abuse researchers consider what possible biases may be generated within the sample and seek to employ sampling techniques that will limit these. In the analysis of data from such studies it is important to compare the findings with work from other studies (Hartnoll, 1997; Stimson et al., 1997) and what is known about the target population, and to remember that caution is needed when trying to infer from the study to the population as a whole (e.g., Topp et al., 2004).

6. MAKING THE MOST OF INCOMPLETE LISTS OF DRUG ABUSERS

The construction of a sampling frame is dependent on having, or being able to construct, a list covering all the members of the target population. In the drug epidemiology field only lists that include problem drug users are readily available but none of these are in any way complete. The type of list typically used is generally drug treatment admissions or discharges, but other lists are available in many countries such as arrests that either identify the arrestee as a drug user or that indicate that the arrest was made for a drug use offence; rosters of people attending social service agencies who are identified as drug users; records from accident and emergency hospital clinics; records from infectious diseases clinics that may identify drug injectors; and, deaths registers that indicate drug-related deaths. There are many limitations associated with use of these lists as are outlined in Chapter 5.

Cross-sectional and longitudinal surveys of heroin users and injecting drug users show that substantial proportions enter treatment, use harm reduction services, get arrested and often imprisoned. For example, a survey of recent injecting drug users in London reported that about 80 percent had attended a syringe exchange; over half had been tested for HIV or HCV; 40 percent had been in treatment; and 20 percent had been in prison since they started injecting. However, it has been found that over the duration of the studies, the proportion of time spent in treatment or in prison or having any contact with any service may be small (e.g., Hser et al., 1992) and there are examples of subjects initiating and ceasing injecting without having experienced any service contact at all (McKeganey and Platt, 1993; Inciardi and Harrison, 2000).

6.1. Specialist Prevalence Estimation Methods

Under-ascertainment in each of many lists is in fact the basis for indirect methods of prevalence estimation. Modeling the under-ascertainment in each data

source and in combinations of data sources and making restrictive assumptions, allows under-ascertainment factors to be calculated for the various sources separately and together. These techniques are described under the headings of Capture-recapture and Multiplier methods, covered in Chapter 8. They are however useful only in estimating prevalence of problem drug use and even then, with some considerable error potential. The error potential lies in two principal areas: the methods are dependent on randomly sampling the drug abuser population or at least approximating this, and the adjustment factors; i.e., the ratio of the unknown population of drug users to the number actually known, are often large.

6.2. Re-Defining the Study Population

Drug abuse researchers have adopted a simple solution to the problem of incomplete ascertainment of the study population by re-defining the study population into one that *can* be sampled and from which statistical inferences can made only to this population and not the population as a whole. Thus a study interested in the behavior of drug injectors might draw a sample of those attending low threshold drug treatment services. Any further inference that is then made from these findings to injectors not in contact with these services would not be supported by statistical procedures but only by "what is reasonable to presume . . . "

6.3. The Where, When and Who of Site Sampling

Sampling people without having a prior exhaustive sample frame but instead taking them from some physical or geographical sample point and if necessary at pre-specified times requires planning. A frequently used site sampling procedure in epidemiological studies is to sample at treatment centers, usually hospitals, for members of the target population. In more difficult circumstances an existing list may not be available and drug abuse epidemiologists must construct one as they go along. Many of the services that work with the most hard-to-reach drug users make it a policy not to collect information that identifies their clients whose attendance may be sporadic and brief. Sometimes samples are collected from known areas where drug users congregate and the task is to develop a sampling frame that reflects those drug users who visit this geographical space. TenHouten et al. (1971) discuss the idea of extending such site samples to a plurality of sites in order to sample more effectively from the whole population under study. Hendricks and colleagues (1992) describe an approach that sketches out a map of the city's likely congregation centers for cocaine use in order to draw a multi-site sample from these centers. Depending upon the frequency of occurrence at the sample points, sampling may be systematic, say every fifth identified individual; randomly selecting one fifth of the individuals in true binomial sampling fashion;

or by exhaustive sampling of all individuals within a given time span. Whichever the precise details, the selection procedure within a site needs controlled, careful execution using well-defined principles in the same way as when using a standard sampling frame.

The most common example in the drug abuse research literature of this sort of sampling approach is in treatment studies where, for example daytime attendees at a clinic, hospital or other center are recruited for the study and handled as a randomly sampled selection from the whole population, potential and actual. Note that in this example the population being studied may have been re-defined deliberately to accommodate the sampling procedure.

6.3.1. Site Coverage, Site Attendances and Weighted Sampling

There are two principal questions concerning site sampling: to what extent is the entire target population encompassed and how often do different individuals attend the site or sites. Drawing a random sample of sites is important in order to be able to make subsequent inferences from the sample to population. Any potential sampling scheme should be considered using some specified conceptual typing such as stratification by physical location so that randomization procedures can be applied within each stratum. Failure to achieve a random sample of sites means that any statistical inferences are restricted solely to the sites actually observed without the ability to generalize to the populations at other sites.

The second important issue related to site sampling that has received little attention in the literature is that unless all population members attend the site(s) in question with the same frequency, simple random sampling procedures cannot be used. If some types of drug abusers, attend twice as frequently as others, then in any given time-period of sampling that group will be twice as likely to be drawn into the sample. Site sampling over-samples subjects in direct proportion to the natural frequency of attendance at the site and it is therefore paramount that in site sampling every effort is made to identify the frequency of attendance of sample members at all the various sites that are being used in the sampling scheme and their weighting applied in subsequent analyses. For example, a study of injectors attending pharmacy-based needle and syringe provision in London required a sampling strategy accounting for the variation not only in the number of clients seen across pharmacies but also in the frequency by which injectors accessed them; some would attend daily picking up a small amount of supplies while others would appear only sporadically but take away enough supplies to last them for longer periods of time (Clarke et al., 2001). Failure to allow for this differential weighting will result in the analysis and description being made of *attendance* and not *attendees* at the clinic or site.

7. MOVING INTO THE UNKNOWN—THE ROLE OF NOMINATION METHODS AND CHAIN REFERRALS

The concept of 'hard-to-access' populations has lead drug researchers to use a number of methods to move the immediate knowledge base beyond those drug abusers observed to the wider drug abusing population. In so doing, the intention is to use any available partial listing or dynamically constructed listing of drug abusers to give wider access to the target population. The two primary methods are both dependent on what are called 'nomination' techniques.

7.1. Nomination Methods

Nomination methods are general estimation methods based on information that individuals in a sample provide about their network of acquaintances (e.g. Biernacki and Waldorf, 1981; Morrison 1988; Stimson et al. 1997). One instance of the application of nomination methods is in estimating drug abuse prevalence using benchmark/ratio methods that require an estimation of the proportion of drug users who are, for example, in treatment (e.g., Taylor, 1997). Simply put, a core sample of drug users recruited in an accessible setting is asked to name (nominate) their drug-using acquaintances and then to say whether these acquaintances have been in touch with any drug treatment centers, health services, or any other similar body within a stipulated time period. It is important when asking these questions that definitions of 'drug user' and 'treatment' are precise and that other qualifiers such as time and geographical location are clearly understood. From this information, the proportion of drug users in treatment can be calculated. This is an obvious adaptation from the more direct question: "What proportion of your drug-using acquaintances have been in treatment?" which is less useful as it is a more difficult question to answer and the answer itself may be subject to rounding off than when asked in two stages. Also, as the baseline number of drug users is not known, weighting each respondent's answer by its statistical reliability is not possible. Clearly there is a trade-off between having information on a greater number of drug users using the nomination technique at the expense of information accuracy. The respondents' personal impressions of the drug-using population and treatment are actually being recorded and the accuracy of the information will depend upon how well the respondent answers or can be encouraged to answer the two questions. It would be inadvisable to ask questions such as the precise frequency of poly-drug use or being tested for HIV as this information may not be known as reliably by respondents other than for themselves.

Nonetheless, the value of nomination methods is that they give the researcher the ability to extend information beyond the core sample of observed drug users to the unobserved population. Although the information so acquired is limited, it can be used to give indirect information on several aspects of drug users' behaviors.

From a sampling point of view, representativeness depends on the selection of the initial sample and, for reasons we have already discussed, this means at best that it is representative of some more restrictively defined population. Theoretical assessments of these methods have been slow but they broadly follow the tradition of two-stage cluster sampling and closely parallel 'star sampling' used in ecology (Thompson and Seber, 1996). Parker and colleagues (1987) offer one of the few comparisons of the impact of different nomination techniques in a study estimating the number of opiate users in four English towns.

7.2. Following-Up the Nominations—The Chain Referral or Snowball Sample

A snowball or chain referral sample is an extension of nomination methods where the initial sample's (termed 'zero stage sample') nominees are traced, interviewed and in turn asked for further nominees. This process is then repeated for a number of further stages (e.g., Stimson et al., 1997). Although a snowball sample can begin at stage zero with sampling from an incomplete list of drug users, it is better to begin with some random site selection procedure. The purpose of this technique is to penetrate into networks of drug users that would be difficult and costly to reach by any other procedure. The value of the approach is confirmed by the long list of informative qualitative and quantitative studies that have used this method for sample generation. Drug use is a socially mediated behavior and drug users as a rule inhabit linked social networks that the snowball sampling procedure exploits.

7.2.1. Snowballing Theory and Snowballing Practice

The following much-quoted key areas for snowball sampling were originally put forward by Biernacki and Waldorf (1981): i) finding respondents and starting referral chains; ii) verifying the eligibility of potential respondents; iii) engaging respondents as research assistants, iv) controlling the types of chains and number of cases, and, v) checking and monitoring referral chains and data quality. It is important to note that only two of these are concerned with the statistical theoretical aspects of snowball sampling and that the remainder deal with practical ground level considerations. This is reflected in the fact that many studies that use this approach have to a large extent ignored statistical sampling issues and have tended to be content with generating a large number of people who exhibit specific target behavior such as drug injecting. Exploratory studies use snowball sampling to penetrate networks of drug users as there are few alternative methods available.

Rapoport (1979; 1980) among others has theoretically reviewed the idea of penetration into social networks. The concerns then shift from whether the sample is randomly drawn to whether representativeness in some sense can be achieved through other controls and to what extent the link tracing can give access to the

whole of a specific network. Mathematical analysis of what are called 'biased networks' has managed to progress only through making unrealistic assumptions that give the crudest practical guide to the researcher. In practice, different link-tracing rules, different drugs, and different social strata, all create different problems that need to be addressed in a research design in the absence of detailed, theoretical guidance from the probability sampling perspective. It would seem possible, for example, that heroin-using networks might be more closely knit than cocaine-using networks, that middle-class drug users might be less prepared to nominate drug-using acquaintances, and that treatment centers will give better access to opiate users than to cocaine users. In these circumstances the only consensus advice appears to be that the initial sample is at least spread as widely as possible across the target population and is as large as possible. It would also seem useful to pursue chains to their full length (sampling to extinction) wherever this is possible.

Snowball samples of this sort can be useful in a number of ways (e.g., Trotter and Medina-Mora, 2000; Fountain et al. 2000). Determining the range of types of behavior patterns or types of individuals is possible by using classification (or cluster) analysis methods that do not need to account for the frequency with which the behavior types occur. Relationships between variables are sometimes more robust than is prevalence to non-representativeness. Modeling the relationship of a particular response (for instance the presence of HIV) to other variables (such as age or injection practices) does not necessarily require accurate representation on these explanatory variables; it requires a sample that does not preferentially contact say HIV cases among young groups. How far a non-random sample can be nearly-representative is influenced by a great variety of factors, amongst them the number of initial recruits, the extent to which contacts form into closed cliques, and the extent to which cliques are connected. Generalization from such a sample to the whole population must be made without the benefit of statistical confidence calculations.

But not all studies have taken this restrictive approach. Others have attempted to build random selection procedures into their sampling, although the theoretical basis for the procedures is often unclear (Bieleman et al., 1993). Goodman (1961) has developed mathematical models that provide statistically valid estimates for the population from a chain referral sample. In practice this means the potential benefits that can come from adopting procedures that statistically underpin the sample have to be balanced against those that generate a large sample and can be accomplished in practice.

That said, many of the methodological and practical difficulties of generating a snowball sample are unrelated to statistical issues. For example, interviewer selection is extremely important, as most successful snowball sampling designs require the contact and interview of drug abusers in naturalistic settings. Some studies have taken this issue further by using members of the target community as interviewers with apparent success in generating sample numbers, particularly

groups that are in some respect closed to outsiders such as drug users from ethnic minority populations. As with all referral techniques, attention to ethical and health and safety issues is of paramount importance. The following references discuss these methods in more depth; Fountain, 2004, Griffiths et al., 1993, and Power and Harkinson, 1993.

7.2.2. Using Chain Referrals in a Sampling Procedure

From a sampling point of view the purpose of the study; be it exploratory, descriptive and modeling or attempting to draw inference about the population, will determine the way the snowball sample is constructed. In defining a chain referral sampling procedure three factors are important: the method for recruiting the zero wave sample, the number of waves (link tracing procedures) the sample uses, and the rules that are used to make nominations.

7.2.2.1. Recruitment. There are no formal theoretical statistical guidelines on what type of zero wave procedure to adopt save only that if a snowball sample is to have any technical statistical validity in making inference to a larger population it must begin with a seeding sample that is randomly drawn and as large as possible. This can be difficult to achieve in practice although the introduction of site sampling in the absence of a population sampling frame has enabled snowball sampling theory to develop along classical statistical lines. Using this combined sampling model has been a major step forward in theoretical terms. If the initiating site sample can be constructed as a probability sample, statistical inferences can be made to the population from both the zero stage sample and from the complete snowball sample. TenHouten et al. (1971) incorporated the two sampling methods, site sampling and snowball sampling, in their analysis of a community's leadership structure.

Snowball samples even if not generated by an initial random sample may still be useful for some other purposes. In these circumstances it is probably advisable to make sure the initial sample is at least spread across the conceived target population as much as possible and it is often recommended that in this first wave all available nominee links be traced.

7.2.2.2. Waves. There is a variety of ways to determine the number of waves of nominee follow-up in a snowball sample. The first and foremost issue is whether they are fixed to be the same for every chain generated by the sample or are allowed to vary naturally as nominations proceed. In spite of ground-breaking work by Goodman (1961), the theoretical consequences even for approximating statistical inferences about population parameters are not clear particularly in terms of efficient use of a given sample size or collection effort or in terms of the degree to which the resulting estimates are unbiased.

For most purposes it would seem inefficient to fix the number of stages in advance. Usually the primary consideration is to obtain enough people by the procedure to allow reasonable numbers for analysis or description. It seems preferable for efficient administrative effort to let the nomination process proceed without imposing constraints on its format. The procedure can be terminated either by sampling to extinction; i.e., until the last wave adds no new nominees or by reaching a pre-determined number of sample members.

7.2.2.3. Nomination Rules. In terms of practical fieldwork, organizing the number of links for selecting nominees at each stage in a snowball procedure is easiest when a uniform set of rules is used for each sample member rather than allowing them to vary haphazardly as the sample waves progress. For any given set of rules, however, Hendricks et al. (1992) note that the theoretical implications are unclear and the best choice might depend rather upon the goal of the study. Each study has different purposes and has both advantages and disadvantages in terms of rate of growth of the sample, costs, and risks of a restricted or biased sample. They provide speculative guidelines which evidence good common sense; for example, tracing all links nominated by the respondent should allow for rapid growth, whereas, tracing links drawn at random may help reduce some concerns about potential biases by maintaining a constant randomization factor in each nomination step. Some work has looked at tracing links in a pre-determined order of importance from the list of the respondent's nominees, taking either a fixed number from or a fixed proportion of the list. For example use of "*first best friend*" and "*second best friend*" and so on may allow for lines of closest or more distant contact to be followed up. Tracing links in a pre-determined reverse importance order (or "*least well known*", "*second* least *well known*" etc) from the list of the respondent's nominees may also produce a well-balanced sample. Although this is perhaps difficult to implement in practice, Rapoport observed (1979) that a tracing through "loose ties" between nominators and nominees will yield a spread sample similar to that derived from a completely randomized network model.

8. CONCLUDING REMARKS AND RECOMMENDATIONS FOR GOOD PRACTICE

As any quantitative drug abuse researcher will point out, talking to even a small number of drug users can produce a wealth of information for understanding behavior that often challenges commonly held beliefs about drug use. For this reason, researchers have been content to dispense with the benefits offered by probability sampling theory and opt for any method that allows them access to sufficient numbers of drug users who exhibit the particular aspect of the behavior of interest. For many purposes the samples that are produced by conventional methods

simply do not produce sufficient numbers for any meaningful analysis or may even exclude those who are often of most interest. Drug studies are sometimes criticized by those naïve to the issues of this area for their lack of sampling rigor, when often in practice less insight would have been achieved by adopting a conventional approach to sample generation.

That said, drug abuse epidemiology has attempted to exploit the analytical benefits from sampling theory, the major driving forces being the attempt to deal with the absence of a sampling frame, working with ambiguous definitions, and encompassing a broad range of behaviors and a correspondingly broad range of research needs. As far as possible conventional survey techniques have been adapted to take into account difficulties of accessing drug users and important study groups such as the school population. But all research studies are faced with the fact that to date all statistically valid estimation techniques require an initial random sample of the hidden population that has been defined as the target population. All the methods discussed above are framed by this requirement and are based on attempting to define the study population in some way that allows a valid sampling frame to be produced or at least approximated.

The easiest way to remain within the boundaries of conventional sampling approaches, at least in theory, is to define the population in such a way that a list can be produced, such as drug users attending treatment facilities, needle exchanges, or other low threshold services. Valuable information has been generated by such studies but statistical inference about drug users not in contact with these services is not possible.

Site sampling approaches, i.e., randomly sampling individuals within randomly selected sites and attempting to correct for variation in frequency of appearance at these sites, provides a solution to the problem if properly carried out. Site sampling can also help supply initial random sampling of the hidden population required by nomination methods and chain referral sampling, where a larger, better initial stage mitigates the clustering effects likely in later stages. It should be noted though that this implies defining the target population as the drug users who will inhabit the space the sites cover and drug users who do not are still excluded. Although they represent an import step forward in theory, it is both practically and methodologically difficult to construct such samples and properly drawn random samples of the general drug-abusing population, which are hidden or are otherwise inaccessible, are simply often not achievable in practice. Other methods are by default the only avenues of potential development in many areas with snowball techniques being the only practical way of generating a sample from many important populations. Furthermore, if nearly-representative samples can be achieved they can be useful in determining the range and pattern of relationships in hidden populations through classification analysis and modeling techniques.

Not all studies require representativeness of the whole drug using population or even the target population. Analysis of within sample differences can be

useful and important policy inferences may not necessarily require statistical underpinning, even if desirable. For example, if a large chain referral sample of drug injectors shows that many are homeless and that high risk injecting practices are commonplace, then this information can be sufficient for prompting policy responses even if it is not possible to statistically quantify these behaviors in the total injecting population. Obviously, such analysis must be made with caution and one would want to be assured that recruitment methods avoided biases wherever possible, especially biases toward homelessness or high-risk behavior. In conclusion then drug researchers are faced with a complex set of practical, methodological and theoretical issues to balance in their research designs. Only some of these are related to sampling but the benefits that probabilistic sampling procedures can bring are sometimes simply unobtainable or have a restrictive impact on the topics for subsequent analysis. A consideration of sampling issues though is still a critical component in any investigation in this area. Drug abuse researchers need to continue to strive to employ probability methods where they can, make a best approximation where this is impossible, always be mindful of the importance of target population definition, and be aware how in practice this may differ from the intended group of drug users who are of interest. Most importantly, caution and sensitivity to possible biases should be used when making inferences from information that does not have the benefit of being based on probability techniques and collaborative evidence should be sought wherever possible.

REFERENCES

Bieleman, B., Diaz, A., Merlo, G., and Kaplan Ch.D. (1993). *Lines across Europe: nature and extent of cocaine use in Barcelona, Rotterdam and Turin.* Svets and Zeitlinger, Amsterdam.

Biernacki, P. and Waldorf, D. (1981). Snowball sampling: Problems and techniques of chain referral sampling. *Sociological Methods and Research* 10 (2), pp. 141–163.

Buster, M.C.A., van Brussel, G.H.A., and van de Brink, W. (2001). Estimating the number of opiate users in Amsterdam by capture-recapture: The importance of case definition. *European Journal of Epidemiology* 17, pp. 935–942.

Clarke, K., Sheridan, J., Griffiths, P., Noble, N., Williamson, S., and Taylor. C. (2001). Pharmacy needle exchange: do clients and community pharmacists have matching perceptions? *Pharmaceutical Journal* 266, pp. 553–556.

Cottler et al., 2001. L. Cottler, S.B. Womack, W.M. Compton and A. Ben-Abdallah, Ecstasy abuse and dependence among adolescents and young adults: applicability and reliability of DSM-IV criteria. *Hum. Psychopharmacol.: Clin. Exp*. 16 (2001), pp. 599–606.

Currie, C., Roberts, C., Morgan, A. et. al. (2004). *Young People's Health in Context. Health Behaviour in School Aged Children (HBSC) Study, International Report From The 2001/2002 Survey, Health Policy for Children and Adolescents*, No 4, World Health Organization, Geneva.

Degenhardt et al., in press. Degenhardt, L., Barker, B., Topp, L., in press. Ecstasy use in Australia: findings from a general population survey. Addiction.

European Monitoring Centre for Drugs and Drug Addiction. (2003). *Annual Report 2003: The State of the Drug Problem in the European Union and Norway*. European Monitoring Centre for Drugs and Drug Addiction, Lisbon.

Fountain, J. (Ed). (2004). *Young Refugees and Asylum Seekers in Greater London: Vulnerability to Problematic Drug Use*. Greater London Alcohol and Drug Alliance, Greater London Authority, London.

Fountain, J., Hartnoll, H., Olszewski, D., and Vicente J. (2000). (Eds.), *Understanding and Responding to Drug Use: The Role of Qualitative Research.* EMCDDA Scientific Monograph Series No. 4., EMCDDA, Lisbon.

Goodman, L.A. (1961). Snowball sampling. *Annals of Mathematical Statistics* 32, pp. 148–170.

Griffiths, P., Gossop, M., Powis, B., and Strang, J. (1993). Reaching populations of drug users by the use of privileged access interviewers: methodological and practical issues. *Addiction* 88, pp. 1617–1626.

Hartnoll R. (1997). Estimating the Prevalence of Problem Drug use in Europe. In: Stimson G.V., Hickman M., Quirk A., Frischer M. and Taylor C. (Eds.), *EMCDDA Scientific Monograph Series (No. 1).* Office for Official Publications of the European Communities, Luxemburg.

Hendricks, V.M., Blanken, P., and Adriaans, N.F.P. (1992). *Snowball Sampling: A Pilot Study on Cocaine Use*. IVO, Rotterdam.

Hibell, B., Andersson, B., Ahlstrom, S. et al. (2000). *The 1999 ESPAD Report: The European School Survey Project on Alcohol and Other Drugs.* The Swedish Council for Information on Alcohol and other drugs (CAN) and Council of Europe Pompidou Group.

Hser, Y., Anglin, M.D., Wickens, T.D., Brecht, M.L., and Homer J. (1992). *Techniques for the Estimation of Illicit Drug User Prevalence: An Overview of Relevant Issues*. National Institute of Justice, Washington, D.C.

Inciardi, J. and Harrison, L. (Eds.) (2000), *Harm Reduction: National and International Perspectives*. Sage Publications, Thousand Oaks, CA.

Johnston, L.D., O'Malley, P.M., Bachman, J.G., and Schulenberg, J.E. (2004). *Monitoring The Future National Survey Results on Drug Use, 1975–2003. Volume I: Secondary School Students.* National Institute on Drug Abuse, Bethesda, MD.

McKeganey N. Barnard M. Leyland A. Coote I., and Follet E. (1992). Female street-working prostitution and HIV infection in Glasgow. *British Medical Journal* 305, pp. 801–804.

McKeganey, N. and Platt, S. (1993). Estimating the population prevalence of injection drug use and infection with human immunodeficiency virus among injection drug users in Glasgow, Scotland. *American Journal of Epidemiology*, 138 (3), pp. 170–181.

Parker, H., Newcombe, R., and Bakx K. (1987). The new heroin users: prevalence and characteristics in Wirral, Merseyside. *British Journal of Addiction* 82, p. 4757.

Power, R. and Harkinson, S. (1993). Accessing hidden populations: A survey of indigenous interviewers. In Davies, P., Hart, G., and Aggleton, P. (Eds.), *Social aspects of AIDS.* Falmer Press, New York. pp. 109–119.

Rapoport, A. (1979). Some problems relating to randomly constructed biased networks. In Holland, P.W. and Leinhardt, S. (Eds.), *Perspectives on Social Network Research.* Academic Press, New York. pp. 119–136.

Rapoport, A. (1980). A probabilistic approach to networks. *Social Networks* 2, pp. 1–18.

Stimson, G.V., Hickman, M., Quirk, A., Frischer, M., and Taylor C. (Eds.), (1997). *Estimating the Prevalence of Problem Drug use in Europe. EMCDDA Scientific Monograph Series (No. 1).* Office for Official Publications of the European Communities, Luxemburg.

Taylor, C. (1997). Introduction to multiplier methods. In: Stimson, G., Hickman M., Quirk A. and Frischer M. (Eds.), *Estimating the Prevalence of Drug Misuse in Europe*. Council of Europe, Strasbourg, pp. 111–112.

TenHouten, W.D., Stern, J., and TenHouten D. (1971). Political Leadership in Poor Communities: Applications of Two Sampling Methodologies. In: Orleans, P. and Ellis Jr., W.R. *Race, Change and Urban Society*, Vol 5, Urban Affairs Annual Review, Sage Publications, Beverly Hills, CA.

Thompson, S.K. And Seber, G.A.F. (1996). *Adaptive Sampling*. John Wiley, New York.

Topp L., Degenhardt L., and Barker B. (2004). The external validity of results derived from ecstasy users recruited using purposive sampling strategies. Drug and Alcohol Dependence, Volume 73, Issue 1, 7 January 2004, pp. 33–40m in Science Direct.

Trotter, R. and Medina-Mora, M-E. (2000). In: *Guide to Drug Abuse Epidemiology.* WHO/MSD/MSB 00.3, World Health Organization, Geneva.

United Nations Office of Drugs and Crime. (2003). *Global Assessment Programme on Drug Abuse (GAP) Toolkit Module 3 Conducting School Surveys on Drug Abuse.* United Nations Office on Drugs and Crime, Vienna.

Vogt, W.P. (1999). *Dictionary of Statistics and Methodology: A Non-Technical Guide for the Social Sciences.* Sage Publications, Thousand Oaks, CA.

Winstock, A.R., Griffiths, P., and Stewart, D. (2001). Drugs and the dance music scene: A survey of current drug use patterns among a sample of dance music enthusiasts in the UK. *Drug and Alcohol Dependence* 64(1), pp. 9–17.