Techniques for the Estimation of Illicit Drug-use Prevalence: An Overview of Relevant Issues

Yih-Ing Hser[1], M. Douglas Anglin[1], Thomas D. Wickens[2], Mary-Lynn Brecht[3] and Jack Homer[4]

Note: The authors constitute a team working on a Cocaine Prevalence Estimation project sponsored by the National Institute of Justice (87-IJ-CX-0042) and are members of the UCLA Drug Abuse Research Group. The project is lead by Dr. Anglin as Principal Investigator, Dr. Hser as Project Director, and Drs. Wickens, Brecht, and Homer, who are experts for the individual methodologies under investigation. Project products are a team effort. Dr. Hser is also supported in part by a Research Scientist Development Award (K01DA00139) from the National Institute on Drug Abuse. Points of view or opinions in this document are those of the authors and do not necessarily represent the official positions of the funding agencies.

---

1 Neuropsychiatric Institute, University of California, Los Angeles

2 Department of Psychology, University of California, Los Angeles

3 School of Nursing, University of California, Los Angeles

4 Department of Systems Management, University of Southern California

Abstract

Issues related to prevalence estimation techniques and their application to determine numbers of illicit drug users are described and reviewed. Major issues include adequate definition of populations, availability and suitability of existing indicator data, and the utility and applicability of selected estimation techniques. Discussion of these topics and suggestions for improving estimation are presented. Given the current level of knowledge of drug consumption patterns and the available indicators, a multiple-mode approach may improve estimation efforts and meet various estimation needs. The conclusions derived for drug use prevalence estimation are also applicable to other content areas such as alcoholism, homelessness, or mental illness.

Key words: Prevalence estimation
Illicit drug use
Hidden population
System dynamics model
Synthetic estimation
Multiple-capture census
Closed-population model
Open-population model

1. Background and significance

Illicit drug use is widespread throughout the United States and the use of certain drugs such as crack cocaine is epidemic within some groups (1). Because of continuing high levels of use and the harmful health and social consequences, drug issues have become an important part of our public policy concerns, affecting everything from international diplomacy to elementary school education (2). Despite this concern, valid estimates of drug-use prevalence are not readily available, nor is there a predominant estimation technique that is suitable in all cases. Knowledge of prevalence levels for different types of drugs is important for policy formulation and implementation, especially in terms of law enforcement strategies, adequate provision of treatment, and suitable targeting of prevention programs. Furthermore, accurate estimates may help to reveal patterns of use that have implications for studying the etiology of drug use or evaluating interventions.

The development and application of improved techniques to provide prevalence estimates is also of considerable interest in the scientific community. Researchers that need reliable estimates include, for example, biologists who estimate wildlife animal populations, criminologists who estimate criminal populations and levels of activity, and public health analysts who are interested in certain disease-affected or mentally ill populations. Recent social concern with human immunodeficiency virus (HIV) disease has focused attention on estimation of the number of individuals, including intravenous drug users, who are at risk for HIV infection. While specific populations of interest vary across disciplines, there are generic aspects of prevalence modeling applicable to all content areas.

The purpose of this paper is to review issues pertinent to estimating illicit drug-use prevalence in the United States. First, we briefly review major common problem areas in prevalence estimation. Second, we explore issues that restrict the utility of various existing approaches, including definition of the target population, use of the existing indicators, and the suitability of various estimation techniques. Finally, we discuss possible approaches that may improve prevalence estimation.

Statement of problem

The task of accurately estimating illicit drug-use prevalence, as aptly described by Boruch (3), is "counting the hard to count and measuring the hard to measure." Because of strong legal and social reasons for non-disclosure of illicit drug-use behavior (4, 5), straightforward methods such as a complete enumeration or probability sampling are impossible. Sampling bias due to discounting the "invisible" or "hidden" components of drug using population is a major problem in prevalence estimates based on conventional sample survey studies. Another source of potential bias is the method of obtaining information on drug-use behavior. Most surveys rely on self-report which often introduces error. The inability to adjust for these biases results in questionable estimates.

As alternative approaches for the estimation of unknown population sizes, researchers have developed and applied many techniques in the area of drug use (6-9), but results have not been entirely satisfactory (10). The published estimates are usually accompanied by wide measurement error ranges or are only applicable to groups of restricted characteristics and lack generalizability to larger populations. Major applications in drug use have used nonsurvey data--for example, drug treatment admissions, drug-related emergency room visits, and drug related deaths. None of these indicators are

probability samples and possibly represent overlapping portions of the drug using population. Such data systems were developed primarily to monitor trends in drug use over time, not for prevalence estimation. Estimation methods applied to these data have often combined several indicators with differential weighting as a partial control for overlap. Other efforts have focused on only one indicator with multiple observations of the same reporting cohort so that the unobservable size could be estimated. However, estimation efforts are still compromised by incomplete or missing information in existing indicators or by the limited coverage inherent in the biased sampling methods currently used. In addition, data inclusion criteria may vary so much among indicators that very different definitions of the target population may be mandated by different indicators. This biased sampling restricts the generalizability of a model's estimates as well as the comparability between estimates resulting from different methods.

Recently, more attention has been paid to the inadequacies of estimation efforts because the spread of HIV has raised considerable concern for estimating numbers of intravenous and other drug users who put themselves at risk for infection by sharing contaminated needles or by other high-risk behavior. In this regard, most of our estimation examples will be based on heroin and cocaine use because these drugs are commonly abused and are often used by injection. Data about the nature and extent of HIV infection among intravenous drug users and other drug users are sparse and considerable data collection efforts have been initiated in an attempt to provide information necessary for policy decisions to slow the spread of HIV. As is the case with drug use prevalence estimates generally, the accuracy of the derived estimates using HIV seroprevalence data is questionable because of large measurement errors and inconsistency of estimates across studies. The published estimates

of intravenous drug users are judged by some to be off by a factor of two in either direction (11). Sampling bias and insufficient information on the base population size are again among the difficulties encountered in these estimation tasks.

Our review of past efforts in prevalence estimation has identified at least four related and complex areas that must be considered to improve estimates: definitions that unambiguously specify the user population; available data for the desired drug types, geographic areas or time periods; data suitability for the estimation method; and estimation technique validity. The remainder of this paper discusses these areas in more detail.

2. Population Inclusion and Classification Considerations

One primary issue with which any prevalence estimation effort must deal is what to measure. In the context of drug use, prevalence can refer to people, events, or amount of drug consumption. In the present paper, we are mostly concerned with counting people. Specifications for prevalence estimation also have to consider time frame and geographic boundaries. This is because drug use is a dynamic process; new users move into the actively using population while current users may cease use due to treatment, incarceration, or death. Drug users also move in and out living areas. Specifications for time intervals or geographic areas are usually pragmatically determined and are mostly straightforward.

Determining inclusions based on past use or current use constitutes temporal criteria of prevalence. Several prevalence categories classified by temporal criteria are common in the study of drug use: lifetime, point, and period. Lifetime prevalence is the proportion of individuals who have ever used a drug. Lifetime prevalence is important for assessing the cumulative impact of drug use on society. Point prevalence is the proportion of drug-

using individuals in a population at a given point in time (the point is usually defined as a 24-hour period prior to data collection). Period prevalence is the proportion of the population using the specified drug during a specific time period. An important period category is current users, who are usually defined as those using in the 30 days prior to data collection. The size of the current-users subpopulation indicates the extent of the immediate drug use problem in an area. Another temporal category of particular interest to government officials is calendar-year prevalence, a period which includes all cases of use during a one-year period.

Of most concern, and perhaps the most difficult definitional task, is the specification of the nature of the drug-use population because of the variation in drug use levels among users. Three major problems arise: first, the nature of the drug and consequences of its use, especially the type or level of use that is of concern; second, the definition of categories, such as "addicts," "occasional users," and so on; and third, the practical problem of obtaining the appropriate sample. The following presents a brief review of the current knowledge of drug-use patterns and consequences.

Specification of the drug of interest is necessary to identify data sources and produce valid prevalence estimates. But drug users often do not confine themselves to the use of just one type of drug. Many studies have shown that concurrent or sequential multiple drug use is common (12). The most popular example is the use of a drug in combination with alcohol. Use of marijuana is also common in combination with use of "hard" drugs such as heroin, cocaine, or others. A frequent practice among intravenous heroin users is "speedballing," the injection of heroin and cocaine, or, less commonly, methamphetamine contained in a single dose. Estimates of users need

to take into consideration this overlap to avoid repeated counts due to polydrug use.

For most drugs, patterns of use are usually described by frequency of use. For example, one schema characterizes patterns of use into five categories: experimental, recreational, circumstantial, intensified, and compulsive (13). Adolescents, for example, often experiment with psychoactive drugs of all kinds as they come in contact with such drugs in their peer culture. Recreational use is the most common pattern of use and the main characteristic is self-control of consumption. Most marijuana and powder cocaine users fall into this use category. Circumstantial users generally take drugs only under certain conditions or in a particular context and not at other times. Intensified use involves a regular pattern of use, sometimes even on a daily basis, in amounts that usually do not result in a level of altered consciousness that impairs work or social functioning. Compulsive use is characterized as high-frequency and high-intensity use of relatively long duration, producing some degree of psychological dependency. These five levels provide a suitable classification system under most circumstances. However, whether one should aggregate these categories or expand them into more categories depends on the purpose of the estimation, availability and suitability of data, and whether a particular schema is required for the estimation model of choice.

All categories of drug users must be considered when prevalence estimation is approached from a view of total drug consumption to which all levels of drug use may contribute. In addition, since individuals may move into and out of various drug-use level categories, being able to monitor the dynamics of such movements is pertinent to the anticipation of changes over time. Information on the rate and duration of drug-use initiation,

maintenance, and cessation, as well as on the process of escalation to greater levels of use, or de-escalation to lower levels, has important implications for developing appropriate models in order to explain drug trends or to anticipate changes in the size of drug-user populations over time. For example, a significant minority (about 25 percent) of heroin users report progressing from first use to daily use in less than one month. Powder cocaine users rarely report this rapid escalation to daily or near daily use. Crack cocaine users, however, have a more rapid escalation to high levels of use that more closely resembles that of heroin addicts.

Because of the associated severe social and health consequences, society is most concerned with "addicts" and with intravenous drug users. Although the term "addicts" is commonly used, disagreement persists among experts as to what constitutes a satisfactory definition of addiction (see, e.g., [14]). Moreover, the detailed information required for theoretical conceptualization (e.g., physiological syndromes) is generally lacking at the individual level in most of the existing large-scale monitoring systems that regularly collect drug use indicator data.

Thus, as a common practice, researchers quite often characterize addiction as compulsive use as defined by high-frequency use, though such use may be different for different drugs. For example, a consistent daily-use pattern may be suitable to describe a heroin addict, but most cocaine dependence follows a pattern of binge use that is more erratic.

Many researchers have come to identify addiction primarily by the consequences of drug use. Past prevalence estimation efforts have relied on indicator data such as drug treatment admissions and drug-related emergency room visits. Such indicator data systems (excepting sample surveys) have been constructed using inclusion criteria that are based on pertinent drug-use

consequences. However, drug-related consequences are often influenced by factors other than drug use itself that may confound estimation results. In this regard, data systems based on consequences may be selectively biased by extraneous influences. For example, public treatment systems are more likely to attract people of lower socioeconomic status, and criminal justice systems typically have a higher concentration of their manpower placed in minority communities. Such differential conditions produce potentially biased data in drug-related indicator systems due to factors other than simple drug use. As a result, the generalizability of estimates based on such data sources may be suspect.

3. Data Available for Prevalence Estimation

Because the availability and suitability of data are important considerations in choosing prevalence estimation approaches and results, this section examines the information currently available from existing drug-use indicator systems.

General information on the extent of drug use in the United States and trends in use can at present be obtained from several sources. Table 1 summarizes the major sources with national scope and comments on their use for prevalence estimation purposes. (Sources of data at the local level are beyond the scope of this paper.) These sources include general surveys and special-purpose federal data systems that contain drug-related information. One major general survey is the National Household Survey on Drug Abuse (15), which has been conducted every two or three years among about 8,000 household residents aged twelve and older. Another important survey data source, the annual High School Senior Survey (16), consists of self-administered questionnaires completed by approximately 17,000 high school seniors. The National Ambulatory Medical Care Survey (NAMCS) (17, 18) gathers information

about patient visits to non-federal, office-based physicians in the nation. The National Hospital Discharge Survey (NHDS) (19) consists of short-stay inpatients discharged from a national sample of non-federal hospitals. Also useful are federal data systems such as the Drug Abuse Warning Network (DAWN) (20) which contains drug-related hospital emergency room visits and medical examiner (corona) mentions, the System to Retrieve Drug Evidence (STRIDE) which provides information on seizures of illicit drugs, the Uniform Crime Report (UCR) database on arrests and convictions for illegal possession or trafficking, the Drug Use Forecasting (DUF) (21) Project which measures drug-use among arrestees as objectively assessed by urine testing, Vital Statistics records on deaths caused by drug use, and the Client-Oriented Data Acquisition Process (CODAP) (22) which contains hospital or clinic records on people seeking treatment for drug-related problems. Each indicator has various strengths and weaknesses. Several of the issues that must be considered in assessing the utility of a data system for prevalence estimation purposes are discussed in the remainder of this section.

3.1 Sampling

Most large-scale surveys of drug use are based on probability sampling techniques. Because of data acquisition considerations and economic reasons, such surveys are seldom based on simple random sampling procedures, with an equal probability of selecting any respondent and independent selection among respondents. However, most of these surveys have chosen sampling designs that achieve representativeness of the target population at the highest level of aggregation. Although these probability samples may not strictly meet assumptions of the prevalence estimation methods, they are commonly used for estimation purposes.

The National Household Survey on Drug Abuse, for example, employs a stratified, multistage sample design based on a predetermined sequence of selection criteria to achieve a national representativeness of households. Individuals included in the survey are selected by using successive sampling units moving from sample locations and households within that location, to the individuals (of specific age, sex, and race) within a specific household that are determined by the sampling plan. Certain subpopulations that are of special interest are often oversampled, but the reported results are appropriately weighted to compensate for unequal probabilities of selection and to reflect the actual distribution of the study population.

General surveys often undersample certain high risk groups because of non-response or non-coverage. Given the difficulties of fieldwork, response rates of about 80 percent achieved by these surveys are a significant accomplishment. Nonetheless, the 20 percent non-response rates are a source of bias in the estimation of drug-use prevalence, especially for low-frequency drug use such as heroin or PCP, where prevalence rates as low as one-half of one percent of the general population are typical. If drug users are disproportionately over-represented among the non-respondents, prevalence rates are substantially underestimated.

Inadequate coverage of high-risk groups is another potential bias. For example, the National Household Survey excludes those in group quarters (military installations, correctional institutions, college dormitories, and hospitals) and those who have no permanent residence (the homeless and residents of single rooms in hotels). Thus, it is likely that the National Household Survey undersamples groups with high rates of use of hard drugs (e.g., heroin or crack cocaine), as well as low-income populations whose members are often transient or cannot afford a household living arrangement.

These groups, according to the mass of empirical literature, are at highest risk for illicit drug abuse (23-25). For example, an analysis of illicit drug use among arrestees indicated that there are two to six times more regular cocaine users in the arrestee population alone than the Household Survey indicated for the whole nation (25). It seems reasonable to treat the NHS results as a lower-bound for prevalence estimation. Other surveys, such as the Epidemiological Catchment Area Survey (ECA) (26), attempted to reduce such selective noncoverage. The ECA included institutional facilities such as mental hospitals and state-operated correctional facilities and nursing homes, but did not include transient facilities such as motels, hotels, dormitories, military installations, or the homeless.

Limitations of sampling that result in underestimation of drug consumption are also exemplified in the High School Senior Survey. Persons who dropped out of school prior to their senior year or students who were absent on the day of the survey were not included in the sample. It is estimated that the dropout rate in the U.S. may average 15 to 20 percent of a birth cohort (27), or even higher in some urban minority populations. It is also known that drug use among dropouts or those frequently absent from school is higher and more extensive for some subgroups than among their peers who continue in school (28). High School Senior Survey staffs' examination of absentees showed levels of drug use consistent with the survey (16). Therefore, estimates of drug use from these data sources are conservative; however, the relative trends in drug use over time are considered reliable.

Other federal indicator systems are not probability samples and may manifest severe coverage limitations. For example, the Drug Abuse Warning Network (DAWN) system, which reports drug-related emergency room visits and medical examiner mentions, covers only non-federal hospitals. The Drug Use

Forecasting (DUF) project, maintained by the National Institute of Justice, monitors drug use by arrestees and places priority in sampling arrestees whose charge is for a non-drug-related offense.

3.2 Validity of Self-Report

The disclosure of stigmatized behavior such as illicit drug use is generally resisted for legal or social reasons. With assured confidentiality and anonymity of their response, the accuracy of self-reported drug use among general population groups is believed quite high (70 to 90 percent) based on checks for internal validity (i.e., estimates of friends' drug use closely parallel cumulative estimates of overall drug use). Some evidence, however, shows that as society has become less tolerant of drugs, people have become less willing to report drug use, even in anonymous surveys. In the High School Senior Survey, 18 percent of the white and 28 percent of the black students stated they may not have reported heroin use if they had ever used heroin (27). Among the arrestee population, only about half the number of arrestees with positive urinalyses self-reported recent drug use (29). Besides deliberate underreporting, memory failure and other cognitive complexities in recalling actual behavior can distort self-reported drug use. Validation of self-report in such general surveys is difficult (30).

3.3 Event-based and Person-based Data Systems

The distinction between event-based and person-based records contained in the available data systems has to be considered in prevalence estimation, since this distinction influences the meaning of the count. Event-based systems are those in which each record arises from a single event, such as an emergency room admission. Person-based systems are those that provide records corresponding to individuals or allow ways to link an individual's multiple

records. Except for survey type data, most federal monitoring systems, such as DAWN or the Client-Oriented Data Acquisition Process (CODAP), are event-based record systems. It is often the case that several records actually belong to one individual who has multiple treatment admissions or emergency room episodes. However, the number of people actually responsible for the number of records in these data cannot easily be determined. For confidentiality and practicality reasons, most data systems do not collect information that allows the identification of individuals. The inability to identify the same individuals contributing to event-based records poses a major difficulty for estimation methods that are based on multiple observations and require such identification.

When one wishes to obtain person-based estimates from an event-based data source, a procedure must be available that provides a unique identifier for matching purposes only and that cannot be used to physically identify a subject. To accomplish this goal, computerized matching techniques have been developed for the CODAP treatment admission records, based on several demographic, treatment and drug use characteristics (31). The method is, however, complicated, difficult to validate, and therefore may not be applicable for widespread use. A relatively unexplored alternative that may be feasible is to obtain respondents' self-reported multiple capture history (e.g., treatment episodes, arrests, emergency room visits, etc.). The problem remains, however, that all the difficulties associated with self-reporting, such as memory failure or under- or over-reporting can bias the resulting estimates.

3.4 Data System Consistency

Trend analysis requires consistency in the reporting panel, a goal which is not often achieved in continuously reporting indicator systems, especially

when agency reporting is voluntary. For example, in 1982 a major change occurred in reporting for the Client-Oriented Data Acquisition Process (CODAP), a nationwide database of federal drug treatment admissions. Prior to that time, reporting was mandatory for all federally funded treatment programs; in 1982 reporting became voluntary and many agencies chose not to do so. However, this indicator system on treatment admissions remains comprehensive and useful for those time periods during which reporting was relatively stable.

Inconsistencies in reporting standards and practices must also be considered. For example, lack of consistency in local reporting systems and regional variations in law enforcement may degrade the usefulness of the Uniform Crime Reports as an isolated indicator for prevalence estimation purposes. Policy changes may also affect the suitability of federal indicator systems in prevalence estimation. For example, law enforcement may shift priority for certain types of arrests. Treatment availability may also change depending on allocated resources. Therefore, interpretation of these indicators is usually not straightforward and needs careful qualification.

Because data systems lack standardized methods of data collection, comparison of indicators can be impaired. The choice of appropriate models or methods must include consideration of the above-mentioned data limitations. We next review the models commonly used in prevalence estimation.

4. Prevalence estimation methods

Whenever the size of a population must be estimated instead of being directly observed, it is necessary to make various assumptions about the phenomenon, the population under investigation, and the observation procedure. Under most circumstances, these assumptions constitute a mathematical or statistical model. Using such models, partial information about the

population is extrapolated to an estimate of the total population, or estimation is made to a new geographic area or time period for which such information is not available.

Several techniques have been applied to the problem of estimating the prevalence of drug use. The discussion below reviews selected prevalence estimation models of varying complexity: synthetic estimation, multiple-capture census, and system dynamics modeling. Several lesser-used models, such as social network analysis (32) and backwards extrapolation (33), are not included in this discussion.

At the start, it is helpful to draw a distinction between static models that describe the sampling process at a single point in time and dynamic models that connect information collected at several observation points. The static models are generally simpler to understand and to apply. However, because they assimilate less information about the drug-use phenomenon, their results may be less valid than those of the more elaborate dynamic models. The dynamic models potentially produce better estimates, but they require many more assumptions about the temporal evolution of drug-use patterns. To the extent that these assumptions are inaccurate, their estimates may fail to a degree much greater than their statistical standard errors imply. Choosing a model that more accurately reflects the complexities of the phenomenon under consideration is intuitively appealing, essential for theoretical development, and often necessary for valid statistical estimation. But there are trade-offs between parsimony and validity. Thus, it is important to consider assumptions of several of the estimation approaches.

An important difference among the models is their ability to identify so-called <u>hidden populations</u>, that is, populations of users who are severely undersampled or completely missing from the data. Clearly the ability to

estimate the sizes of these hidden groups is critical to the prevalence-estimation enterprise, but equally obviously things never observed cannot be measured without making some strong assumptions about the process. In general, the simpler models are less capable of estimating these hidden populations while the more structured models do a better job, although often at risk of misspecification. The same trade-off between parsimony and validity applies here as well.

4.1. Static models

The least complex of the prevalence estimation models are those that describe the system at a single point of time. These models have been the most popular in drug use prevalence estimation.

4.1.1. Synthetic Estimation

The simplest of the static estimation techniques are those that employ synthetic estimation. These methods project prevalence estimates from several more readily available sources or indicators to new populations. The crux of the synthetic estimation methods is the selection of an appropriate set of predictor variables (e.g., ethnicity, gender, age, and regional location) and the determination of the proper weights to be applied to them. The many ways of weight determination that have been used range from simply transferring relationships found in one population to another population (34), to some rationalized linear function (35), to factor analytic modeling (6, 7).

Population projection models were originally developed by the National Center for Health Statistics for obtaining estimates of prevalence in areas such as cities or Standard Metropolitan Statistical Areas (SMSAs) (34, 36). Their logic is that if drug use prevalence rates are known in one population with known demographic distribution, the relationships between prevalence and

demographic characteristics can be transferred to another population, either smaller or larger than the first.

Specifically, suppose that the population can be categorized on the basis of a set of mutually exclusive and exhaustive classes such as age, sex, and race. In a well-studied reference population the proportion of drug-using members of each category is estimated. In a new population, a demographic survey indicates the frequencies in the age-sex-race combinations. Combining this frequency information with the rates from the reference population (as a weighted sum) gives the synthetic estimate. This simple weighting scheme can be modified in various ways; for example, regression methods can be used to include ancillary information in the estimate (34).

The population projection method is essentially data-driven and does not require assumptions about the process and time course of drug use. Thus far, most applications of the population projection technique to the estimation of drug use prevalence have relied on the National Household Survey in combination with census data. This method has been widely used because of its simplicity and because population data and the weight coefficients related to drug use are easily obtainable from the National Household Survey.

The use of synthetic estimates has been questioned for several reasons. A potential problem lies in its total reliance on survey-based prevalence data. Survey data on stigmatized behavior are particularly subject to sampling bias and distorted reporting. Another possible problem is that unmeasured characteristics may make the rates in the target population different from those in the referenced population. For example, the method will not produce valid estimates when regional differences (which are considerable in many indicator data) in drug use and availability render the demographic variables insufficient to yield an adequate estimate. Finally,

the method has no statistical basis, so no confidence intervals for the estimate can be obtained. The direct form of synthetic estimation should only be applied when samples are sufficiently representative so that the observed pattern can be projected from, or back to, another population.

In contrast to the population projection form of synthetic estimation which emphasizes the demographic characteristics of a particular area, the principal components approach uses the relationships observed among multiple indicators in several geographic areas (e.g., SMSAs) in an attempt to obtain a single composite indicator of drug use. There are many indicators that may be related to drug use (see Table 1 for examples). Each of these indicators is subject to measurement and sampling error, but each reflects some aspect of an underlying construct, namely, drug-use prevalence in the area. By combining these indicators with appropriate weighting into a single composite index, one can, in principle, derive an index which is more reliable than any single indicator alone.

The "Heroin Prevalence Index" (HPI) of Person, Retka, and Woodward (6, 7) illustrates the principal components procedure based on rank-orderings of SMSAs by several indicator measures. The three-stage technique involved the calculation of the HPI from a principal components analysis of indicators, its calibration against independent estimates of prevalence in at least two areas used as reference points, and its use to project drug use estimates in new geographic areas (e.g., other SMSAs).

The usefulness of the HPI approach largely depends on the acceptance of its assumptions. Difficulties can arise at each of the three steps. In the first step, the principal component approach requires the assumption that the measures are monotonically related. Except for measurement error, the rank ordering of the sampling units (e.g., SMSAs) on one indicator should be the

same as the rank ordering on the other indicators, and this rank should be the same as the rank ordering on the true underlying prevalence. Possible violations of monotonically rank-ordered relationships have been pointed out by Demaree et al. (37); for example, given limited treatment resources, the probability of admission for a heroin user may decrease the more heroin users there are, or simply because the reporting bases underlying the indicators are different between SMSAs. In addition, unless the variables are standardized, a principal component analysis gives greatest weight to the indicators that have the maximal variance. If this variance is related to some underlying relationship other than true prevalence, then the resultant measures reflect these aspects rather than actual drug use. An example of a potentially overweighted measure is the federal resource availability for treatment.

The second step requires the HPI to be calibrated to match independent estimates of prevalence in two or more SMSAs. This matching can be no better than the quality of these independent estimates. Unfortunately, well-based independent estimates are not always available. Biases in these estimates also affect the quality of the calibration. Moreover, differences in the definition of prevalence used in the two anchoring areas affect the calibration at intermediate values of the HPI.

The third step, in which the estimation is actually made, requires that values of the measures used to determine the HPI are available for the target populations. This requirement usually forces the units for the projected population to be the same as those for the populations used to derive the HPI--for example, SMSAs. Finally, the linear nature of the function linking the HPI to prevalence may be suspect.

In summary, the synthetic-estimation approaches allow projection of estimates for geographic areas lacking such information. Valid estimates

derived from such approaches require, at the least, selection of indicator data that satisfy certain specific properties (e.g., monotonic relationships). Synthetic estimation also requires availability of good quality independent estimates in two or more equivalent geographic areas for reference points.

4.1.2. Closed-population capture models

Another variety of static estimation procedure is closed-population multiple-capture methods, using two or more surveys (or different data sources such as emergency room and arrest records) to probe the population during the same time frame. Each survey must be able to identify individual cases and determine which of them have been detected in each of the overlapping surveys. Using the information about the relative sizes of the samples and their overlaps, one can estimate the number of individuals that have not been detected. The population whose size is to be estimated here is said to be closed because of the single sampling time. This procedure has also been referred as Dual-System estimation (38, 39). The cross-sectional nature of this situation contrasts with the longitudinal open-population models that are discussed below.

As one example of multiple-capture procedures, consider a pair of surveys aimed at detecting the members of a population. The detected individuals are cross-classified in an incomplete 2x2 table containing frequencies of those sampled in the first survey but not the second, those sampled in the second but not the first, and those sampled in both. Only these three cells actually contain data, the frequency of those never observed is unknown and must be estimated. With more than two surveys, a higher-dimensional table is obtained, always with one missing cell for those not detected in any survey. To obtain an estimate of the total prevalence, the frequency in the unobserved cell is estimated.

The estimation of the unobserved frequency is usually accomplished by fitting a simple association model to the incomplete table (40-42). The more samples that are available, the more complex this model can be. In the two-sample case, one must make the very restrictive assumption that observations in the two samples are independent of each other so that frequency of the missing cell can derived by multiplying the two marginal frequencies and divided by the frequency of twice observed cell. With several samples various forms of association among the frequencies can be allowed while the absence of the highest-ordered association must be assumed. This type of model has been used by the Census Bureau for population estimation (39).

The static, closed-population multiple-capture models are conceptually and arithmetically simple. To the extent that the different surveys probe the population in somewhat different ways, they can combine several weaker sources of data into a stronger conclusion. In this sense they provide some hope of a way to extrapolate to poorly-measured populations. They are limited, however, by their rather restrictive data requirements. There must be a series of surveys, and each member of the population must be equivalently at risk for detection by each survey. Static models may be inappropriate when the samples are obtained at different times. In addition, the comprehensiveness in coverage of each survey, the accuracy of matching individuals across surveys, and the adequacy of independence of the surveys are among the difficulties in applying multiple-capture methods in general. However, if data considerations can be satisfied, closed-population models are simple and easy to apply. Confidence in the results can be judged by statistically derived confidence intervals and identified boundaries for generalizability are available because the general statistical properties of such models have been well studied.

4.2. Dynamic models

Dynamic models describe changes in the drug-using state of the individual or system and link observations over time, a procedure which potentially obtains a more accurate estimate of prevalence than static models.

4.2.1 Open population recapture models

The data that are drawn from a system of successive censuses with identifiable individuals in the open-population multiple-capture method are similar to those obtained in the closed-population multiple-capture situation described above. These data also form an incomplete contingency table. The similarity is superficial, however, for the processes to which the models are applicable are quite different. In the closed-population situation, every individual is at risk for every census, while in most longitudinal sets of censuses, some individuals leave the population before the final census while others enter after the first census is complete. Therefore, closed-population models trace and estimate the size of a single population, while open population models keep track of population changes and provide estimates reflecting such time-related changes. The data from a longitudinal dynamic process require a different type of statistical model, based on some assumptions about open populations.

Most applications of the multiple-capture methodology to longitudinally repeated drug-use samples have used closed-population models (e.g., [43-46]). Such estimates are potentially biased because some individuals may not be available throughout the entire sequence of time sampling and the degree of bias may increase with longer intervals between samplings. The nature and magnitude of these biases have not been studied. However, the magnitude of the standard errors in these closed-population applications is usually quite large, ranging from 10 to 80 percent of the estimated population sizes. This

problem is especially serious within geographical areas of small population. Models that accommodate the more realistic open-population dynamics are more appropriate for estimation purposes.

The truncated Poisson estimation model is the simplest version of the multiple-observation models, applied in situations where only the number of times that an individual appears in a data system is recorded, without attempting to cross-tabulate the specific occurrences. The result is a frequency distribution, starting with the count of individuals observed once and continuing upward. The unobserved portion of the population, for which a count is desired, resides in the missing "zero" cell of this distribution (i.e., those never observed). To estimate the size of this cell, one fits an appropriate probability distribution to the balance of the empirical distribution and uses its form to estimate the number of missing observations. When the index event is rare, as it usually is for drug-use incidents, the distribution plausibly has a Poisson form, leading to truncated Poisson estimates (47). To estimate the size of the population, an incomplete Poisson distribution is fit to the frequencies of the observed events, and the single rate parameter of the distribution is estimated. Knowing this parameter, the size of the unobserved category is estimated and added to the observed count to obtain the final estimate.

The truncated Poisson models have been used for estimating the size of the criminal population from arrest history records (48) and the number of persons engaged in drug-related crimes from drug-related arrest data (49). In these applications, an arrest distribution was constructed from the number of observed arrests and the number of arrestees responsible for them, then the truncated Poisson estimation procedure was applied to derive the population estimates. Based on a similar rationale, Research Triangle Institute (50) has

applied the model to estimate the size of the treatment-susceptible heroin population. Their model was unusual in that two sources of data were utilized to estimate the Poisson rate. The number of treatment admissions was available from one data source while the distributional information was derived from a separate, nonlinked source. In this two-source implementation, the comparability of the two populations is critical to ensure that the assumed Poisson distribution is applicable.

The strength of the truncated Poisson method lies in the simplicity of its data requirements and its straightforward statistical formulation. As long as the data can be consolidated into a frequency distribution of the number of people at each level of the repeated observation, an estimate is easily obtained. However, the quality of the estimates depends on the degree to which the Poisson model is an adequate description of the underlying distribution. In particular, the counts must be independent Poisson events. This assumption is frequently violated--for example, criminals are strongly motivated to avoid rearrest and are, to some extent, quite successful in doing so. On the other hand, risk at arrest may increase as the offender becomes known to the police. The effects of violations such as these on the truncated Poisson-derived estimates are unknown.

A more sophisticated approach utilizing repeated sampling may provide a better estimation methodology. Ecological open-population models for the assessment of the size and character of a biological population based on repeated marked samples have been developed and extensively analyzed (for reviews see [51, 52]). The most common class of ecological sampling models that has been suggested as applicable to the estimation of the number of drug users is the capture-recapture type. For an open population, the most recent models are the Jolly-Seber model (51, 53-55) and a related model by Cormack

(56-58). However, a number of these models' assumptions have been judged to be unrealistic for drug-use populations and none of these models has yet been applied to drug-use prevalence estimation.

Another variety of the open-population multiple-capture approach developed recently is a **Markov-based dynamic recapture model** (59). Instead of counting the individual captures in each sample, a longitudinal model is based on the variety of capture histories. The model characterizes capture probabilities by a two-step sampling process. The initial sampling probability is governed by a stochastic process representing users being initially drawn from an infinite population of non-users. After this first sample, the balance of the process is governed by the dynamics of the state structure which represents the evolution of drug-consumption patterns and their repeated observations by some indicator system, e.g. treatment admissions. This process forms a Markov chain (60). The full history probabilities are therefore the product of three terms: (1) the size of sample when an individual is first observed; (2) the probability of that observation; and (3) the probability of the observation history subsequent to the initial observation. This procedure generates estimates of the size of the population from which the observations are drawn.

This Markov estimation technique has the advantage of being able to provide a somewhat more realistic dynamic description of the drug-use process. Its weakness is that the model comes with some strong intrinsic statistical assumptions. In addition, to be practically applied, the model can have only a minimal, and thus limited, dynamic structure if its parameters are to be identifiable and estimated from a sufficiently rich data sources.

Although the open-population models potentially allow a more realistic picture than do closed population models of drug-using populations as they

evolve over time, such models still require certain restrictive assumptions. Some of these assumptions are particular to the models developed in ecological situations and others serve to practically simplify the statistical models so that parameters can be estimated adequately. The latter assumptions include the requirements individuals behave independently and identically with respect to the model's parameters. A number of applications have shown that these assumptions are often violated in biology and health sciences (e.g., [61-63]), and simulation studies have shown potentially large bias in population estimates under such conditions (e.g., [64, 65]).

Theoretically, surveys or data sources used by multiple-capture methods (e.g. open or closed) should be comprehensive in population coverage. In reality, data suitable for this type of application have been limited to treatment admission records that do not provide such comprehensive coverage, thus limiting the generalizability of the estimation results. One common difficulty in applying any multiple-capture models is the necessity of matching individuals across observations. The data source must ensure that an individual captured in one sample is identifiable as the same person if captured in another sample. However, such information is difficult to obtained as discussed in Section 3.

Because the multiple recapture models attempt only to estimate the number of unobserved individuals comparable to the observed sample, they cannot extrapolate to unobserved subpopulations. Thus, they provide no solution to the hidden population problem as it was defined above. However, the probability-theory based open population models generally have better known statistical properties that allow estimates to be evaluated by standard methods such as confidence intervals. Boundaries of generalizability usually can be inferred by data coverage and model specifications.

4.2.2. System Dynamics Modeling

System dynamics is a general methodology used to model systems-level relationships among constituent components, or variables. First developed in the late 1950s, this approach analyzes dynamic phenomenon through feedback-oriented computer modeling. A system dynamics model consists of an interconnected set of difference equations representing continuous-time flows and accumulations of people, materials, and information. After being assigned initial conditions, this set of equations is capable of generating output, over time for each modeled variable, that may represent the true course of events.

System dynamics models typically attempt to explain the observed dynamics of a specified system as being the consequence of endogenous feedback relationships among constituent variables. This endogenous perspective distinguishes the system dynamics approach from other approaches, such as statistical regression, which rely upon exogenous or independent predictors whose own behavior over time is left unexplained by the model. The continuous feedback perspective of system dynamics also leads to models that tend to be larger in scope than Markov-type models because they contain a greater variety of system variables. This enlarged scope is rendered manageable in practice by modeling flows in the aggregate, rather than by keeping a unique record of every individual unit in the flow as is done in the multiple capture models. Because of its "big picture" approach to modeling, system dynamics appears to be a promising technique for tying together a diversity of indicator data and seeking inferences regarding prevalence that are consistent with the multiple data sources.

Examples of system dynamics modeling studies of illicit drug use include the "Persistent Poppy" model of Levin, Roberts and Hirsch (66), models

developed by Gardiner and Shreckengost (9, 67-69), and a recent prevalence model developed by Homer (70). The Persistent Poppy model examined heroin use in New York City from the standpoint of policy rather than prevalence estimation. Although the model contains several interesting endogenous factors--such as law enforcement activity and educational and treatment programs, the model was developed at a time when the numerical data needed for its calibration and validation were lacking. Models developed by Gardiner and Shreckengost addressed the issue of drug supply and demand on a national level with specific application to heroin and cocaine. The central variable is "relative abundance," a comparison of supply with demand, where supply is associated with imports (modeled as exogenous), and demand is determined indirectly by relative abundance itself. The model developed for heroin was additionally used to estimate the number of users, but such estimates were shown to be rather sensitive to the specification of user categories (69). Moreover, even if these categories could be specified appropriately, the model of Gardiner and Shreckengost may lack sufficient feedback structure and internally generated "momentum" to be useful for prevalence estimation and projection.

A significant advantage of the system dynamics approach is its ability to tie together disparate sources of information into an integrated framework for estimation. Thus, it is possible to integrate information about the nature of drug use with incomplete prevalence data to estimate the frequency of unobserved categories and potentially solve the hidden population problem. This use is well-illustrated in a recent national model of cocaine use (70). When appropriately calibrated, this model reproduces historical indicator data (from 1976 onward) and produces prevalence estimates and near-future projections for several population categories, including recreational and

compulsive users. Compulsive user categories are usually not obtainable from general survey data. The estimates of this hidden population requires two sorts of information. First, information is needed regarding the causal or structural relationships that relate the hidden population to the rest of the overall system. For example, the model considers the pattern of escalation from recreational use to compulsive use. Second, numerical data are needed that have a primary logical relationship to the hidden population and that let the model be calibrated appropriately. For example, it is likely that cocaine-related morbidity and mortality are more closely associated with compulsive users than with casual users, and that changes in compulsive cocaine use should therefore resemble patterns seen in the DAWN data. With sufficient numerical data and knowledge of structure, the range of estimates for the hidden population may be narrowed to a considerable degree.

System dynamics is an attractive approach for prevalence estimation largely because it allows for the explicit interlinking of a wide variety of causal factors that drive each other iteratively over time. It can be used to explain past history and to fill in gaps in indicator data. Through the appropriate manipulation of input parameters, it can also be used to project outcomes under different assumed scenarios and policy interventions. However, the very flexibility of system dynamics opens the door to potential model misspecification, a danger which becomes greater as the number of variables increases relative to the quantity of available and relevant data, and as the scope of the model expands to include more diverse phenomena. Like any method, it is sensitive to the quality of the data it uses. Estimates of hidden populations may be particularly sensitive to errors in model specification and the use of inadequate data.

Because of the complexity of system dynamics models and the risks of misspecification, a variety of tests for building confidence in these models have been offered that go well beyond the usual requirement that historical data be reproduced (71). But these validation techniques are themselves subject to uneven application or improper interpretation. It must be recognized that system dynamics modeling, despite its many attractions, is difficult to master and there are many pitfalls in its application that one must be careful to avoid.

4.3 Summary

None of the prevalence estimation methods, reviewed here or known to the field, can provide estimates without knowledgeable and careful application, and without conditional limitations. Some of these limitations are due to the unrealistic nature of the model assumptions, others are due to the demands for specific data of quality. Static models typically take "snapshots" of the drug use problem of an area, and some give additional descriptions in terms of demographic distribution. Dynamic models provide prevalence estimates considering time-related process; they may also specify parameters of the processes and offer some forecasting capability. Synthetic estimation relies on prevalence estimates from independent sources to extrapolate and provide estimates not available for the desired geographical areas. These independent estimates, in most cases, have come from surveys which are subject to numerous criticisms as discussed in Section 3. Multiple-capture types of models provide estimates of an incompletely observed population by projecting from the capture pattern of samples observed over time without requiring independent estimates. Data of sufficient richness to support this type of model application, however, are not readily available. The system dynamics models have the potential to estimate undersampled populations and may provide

a better understanding of the mechanisms and dynamics that may influence the prevalence of drug use and its change over time. However, a system dynamics model is usually built for a specific application purpose and requires considerable specific data, thus restricting its generalizability. Furthermore, the building, calibration, and validation of a system dynamics model are generally difficult and require special caution and expertise.

Strengths and difficulties of these methods are summarized in Table 2. It is clear that methods differ in their data requirements and statistical properties and thus in their utility for being mapped onto the general phenomena of drug use. Notice that the strength of a methodology is quite often characterized as its difficulty as well. For example, although system dynamics modeling provides a larger framework to examine the phenomena and is able to utilize data from many sources, many such models lack sufficient restrictions for the complexity of the structure that they contain. Open-population multiple-capture models focus on a minimal dynamic structure which tends to unrealistically simplify the processes involved, often using, for example, only those data concerned with treatment admissions. However, these models can be statistically tested--a clear strength. Synthetic estimation uses prevalence figures from more general data and is useful because local planning needs must, of necessity, accept estimates made in the absence of good local data. Because independent local surveys are generally expensive and not feasible for local agencies to conduct, synthetic estimation fulfills a clear need. However, the quality of these estimates depends totally on the quality of the larger, more general data, and a direct mapping to a smaller area is not always appropriate.

Different types of data are best suited for one estimation method or another, and estimates derived from these methods must be interpreted within

the appropriate context. Both synthetic estimation and multiple-capture models require person-based data, while system dynamics models can use event-based data. Choices of models have to consider the appropriate use of information that is available. When data representativeness is not an issue, synthetic estimation methods are the least costly and easiest way of providing demographically and geographically adjusted estimates. When multiple observations about identifiable individuals are available, multiple-capture models may be appropriate choices because of their ability to integrate such sequences. If observations are obtained from separate, independent sources and are made during the same time frame, closed-population models are applicable. Open-population models are used to their best advantage to reflect the dynamic aspect of drug-use evolution when individuals are traced over time. Because these multiple-capture models are based on probability theory, their statistical properties are mostly well-known. When appropriately applied, confidence in the estimation results can be judged by conventional confidence intervals and other statistic tests. System dynamics models are applied when multiple sources of information about the system over time are available and when the primary interest is the understanding of the dynamic interrelationships among these indicators within the system. System dynamics modeling is also an excellent tool for making projections and for answering "what-if," or policy simulation, questions.

5. Concluding comments

Although numerous difficulties are associated with prevalence estimation in whatever content domain it is attempted, estimates are necessary in order to make practical decisions for resource allocation, program planning, and other purposes. Recognition and understanding of these difficulties promotes the appropriate use of the available data and provides more defensible

prevalence estimations. Development in at least three areas is needed to improve estimation of drug prevalence:

(1) Improving the understanding of drug-use phenomena.

Information about production, distribution, and consumption characteristics allows choice and specification of models better representing drug use that will lead to better estimates. Continuous monitoring of the phenomenon should lead to a better understanding of the processes and provide a basis for improved prevalence models and improved prevalence estimates.

(2) Consistent, comprehensive and accurate data collection systems.

With a few exceptions, there have been no persistent or consistent efforts to provide continuous data measurements of sufficient coverage that allow statistical techniques to be applied to yield high-quality estimates. The data services have problems with missing many individuals at high risk for drug use and with the questionable validity of self-report. Moreover, prevalence estimation suffers because linkage among the various indicators does not occur. For example, it is not known how many drug users entering treatment have had recent emergency room care for drug related health problems, and vice versa. If such information were available for all indicators, mapping these indicators to provide a non-overlapping prevalence estimate would allow more precise, consistent and functional relationships to be specified.

(3) Continued development and improved utilization of appropriate prevalence estimation techniques.

Current drug-use prevalence estimation is flawed by the deficiencies of the existing data systems and limitations of estimation techniques. Unless major efforts are made to provide a complete count of the user

population--a possibility that seems extremely unlikely--better estimates of the number of individuals using various types of drugs must rely on improving existing or developing new statistical techniques to remedy the data deficiencies. The techniques discussed in Section 4 have all made their contribution to prevalence estimation. However, because of the complex dynamic nature of drug use and the restrictive sampling or incomplete information in the existing indicators, no single method can adequately produce estimates for all categories of users. A single method will never be adequate to meet the heterogeneous needs for different types of prevalence estimation. A variety of complementary methodologies will always be necessary.

It must be realized that an estimate derived by any one particular method is the result of an interplay between theory, methodology, and empirical data. Choices of the estimation model depend on the phenomenon under study as well as on the available data. By applying multiple approaches, each capitalizing on some salient aspect of the prevalence problem, confidence in the results is increased or, at the least, inconsistencies identified. In addition, alternative models using different approaches and data sources are necessary to validate each other when their estimates overlap. The resultant multiple-mode approach is regarded as appropriate in prevalence estimation. Considered together, multiple methods using multiple data sources provide estimation ranges that set boundaries for policy decisions on allocations for enforcement, treatment, and prevention.

References

1. Adams, E. H., and J. Durell (1984). Cocaine: A growing public health problem. In J. Grabowski (Ed.), Cocaine: Pharmacology, Effects, and Treatment of Abuse, National Institute on Drug Abuse Research Monograph 50, pp. 9-14, DHHS Publication No.(ADM)84-1346. Washington, D.C.: U.S. Government Printing Office.

2. Reuter, Peter, Crawford Gordon, and Jonathan Cave (1988). Sealing the Borders: The Effect of Increased Involvement of the Military in the Drug Interdiction Program, R-3594. Santa Monica, California: The RAND Corporation.

3. Boruch, Robert (1988). Counting the Hard to Count and Measuring the Hard to Measure: An Outline of Methods Bearing on Longitudinal Study of Delinquency and Desistance. Paper prepared for Working Group on Desistance, MAC/NIJ Longitudinal Multicohort Study at May 1988 Meeting in Pittsburgh.

4. Blumstein, A., P. C. Sagi, and M. E. Wolfgang (1973). Problems of estimating the number of heroin addicts. In Drug Use in America: Problem in Perspective, Vol.II Social response to drug use, The Technical Papers of the Second Report of the National Commission on Marijuana and Drug Abuse, pp.201-211. Washington, D.C.: U.S. Government Printing Office.

5. Amsel, Z., W. Mandell, L. Matthias, C. Mason, and B. Hocherman (1976). Reliability and validity of self-reported illegal activities and drug use collected from narcotic addicts. The International Journal of the Addictions, 11(2):325-336.

6. Person, P. H., R. L. Retka, and J. A. Woodward (1976). Toward a Heroin Problem Index-an Analytical Model for Drug Abuse Indicators. DHEW Publication No. (ADM)176-367 and (ADM)78-367. Washington, D.C.: U.S. Government Printing Office.

7. Person, P. H., R. L. Retka, and J. A. Woodward (1977). A Method for Estimating Heroin Use Prevalence. DHEW Publication No. 017-024-00589-4. Washington, D.C.: U.S. Government Printing Office.

8. Woodward, J. A., M. L. Brecht, and D. G. Bonnet (1985). Longitudinal Models of Heroin Use. Final report, contract No. DEA-83-16, Drug Enforcement Administration.

9. Shreckengost, R. C. (1985). Estimating cocaine imports into the United States. Unpublished manuscript dated October 1985.

10. Rouse, B. A., N. J. Kozel, and L. G. Richards (Eds.)(1985). Self Report Methods of Estimating Drug Abuse: Meeting Current Challenges to Validity, 181 pp. National Institute on Drug Abuse Research Monograph Series 57. Washington, D.C.: U.S. Government Printing Office.

11. Spencer, Bruce D. (1989). On the accuracy of estimates of numbers of intravenous drug users. In Charles F. Turner, Heather G. Miller, and Lincoln E. Moses (Eds.), AIDS, Sexual Behavior, and IV Drug Use, pp.429-446. Washington, D.C.: National Academy Press.

12. Wesson, D. R., and D. E. Smith (1979). Treatment of the polydrug abuser. In Handbook on Drug Abuse, edited by R. L. DuPont, A. Goldstein, and J. O'Donnell. Washington, D.C.: U.S. Government Printing Office.

13. Siegel, R. K. (1984). Changing patterns of cocaine use: longitudinal observation, consequences, and treatment. In J. Grabowski, Cocaine : Pharmacology, Effects, and Treatment of Abuse, pp. 92-110. National Institute on Drug Abuse Research Monograph 61. Washington, D.C.: U.S. Government Printing Office.

14. Edwards, G., A. Arif, and R. Hodgson (1981). Nomenclature and classification of drug- and alcohol-related problems: a WHO memorandum. Bulletin of the World Health Organization, 59:225-242.

15. National Household Survey on Drug Abuse, Main Findings (1977, 1979, 1982, 1985, and 1988). National Institute on Drug Abuse.

16. Johnston, Lloyd D., Patrick M. O'Malley, and Jerald G. Bachman (1989). Drug Use, Drinking, and Smoking: National Survey Results from High School, College, and Young Adults Populations 1975-1988. Rockville, Maryland: U.S. Department of Health and Human Services.

17. Bryant, E. Earl (1988). Sample Design, Sampling Variance, and Estimation Procedures for the National Ambulatory Medical Care Survey. Hyattsiville, MD: U.S. Dept. of Health and Human Services, Public Health Service, National Center for Health Statistics.

18. Nelson, Cheryl (1988). The National Ambulatory Medical Care Survey: United States, 1975-1981 and 1985 Trends. Hyattsville, MD: U.S. Dept. of Health and Human Services, Public Health Services, Center for Disease Contorl, National Center for Health Statistics.

19. Graham, Dorothy (1988). Detailed Diagnoses and Procedures for Patients Discharged from Short-Stay Hospitals. Hyattville, MD: U.S. Dept. of Health and Human Services, Public Health Services, National Center for Health Statistics.

20. Drug Abuse Warning Network (DAWN). Statistical Series, Annual Data, National Institute on Drug Abuse.

21. Drug Use Forecasting (DUF). National Institute of Justice.

22. Client Oriented Data Acquisition Process (CODAP). National Institute on Drug Abuse.

23. Gandossy, R. P., J. P. Williams, J. Cohen, and H. J. Harwood (1980). Drug and Crime: A Survey and Analysis of the Literature. National Institute of Justice. Washington, D.C.: U.S. Government Printing Office.

24. Robins, L. N., and E. Wish (1977). Childhood deviance as a developmental process: A study of 223 urban black men from birth to 18. Social Forces, 56:448-473.

25. Wish, Eric D. (1990). Drug policy in the united states: Time for a dose of reality. The International Journal of the Addictions.

26. Eaton, W. W., and L. Kessler (1985). Epidemiologic Field Methods in Psychiatry: The NIMH Epidemiologic Catchment Area Program. Orlando, Florida: Academic Press, Inc..

27. Johnston, Lloyd D., Patrick M. O'Malley, and Jerald G. Bachman (1988). Use of Licit and Illicit Drugs by America's High School Students 1975-1984, pp.34-37. Rockville, Maryland: U.S. Department of Health and Human Services.

28. Kandel, D. B., and D. R. Maloff (1983). Commonalities in drug use: A sociological perspective. In P. K. Levison, D. R. Gerstein, and D. R. Maloff (Eds.), Commonalities in Substance and Habitual Behavior. Lexington, Massachusetts: Lexington Books.

29. Harrison, Lana D. (1990). The Validity of Self-Reported Drug Use among Arrestees. Washington, D.C.: National Institute of Justice.

30. Harrell, A. (1985). Validation of self-report: The research record. In B. A. Rouse, N. J. Kozel, and L. G. Richards (Eds.), Self-Report Methods of Estimation Drug Use: Current Challenges to Validity. NIDA Research Monograph 57, DHHS Publication No.(ADM)85-1402. Washington, D.C.: U.S. Government Printing Office.

31. Woodward, J. A., R. L. Retka, and L. Ng (1984). Construct validity of heroin abuse estimators. The International Journal of the Addictions, 19:93-117.

32. Frank, Ove (1979). Estimation of population totals by use of snowball samples. In P. W. Holland and S. Leinhardt (Eds.), Perspectives on Social Network Research, pp. 319-347. Washington, D.C.: National Academy Press.

33. Brookmeyer, Ron, and Mitchell H. Gale (1988). A method for obtaining short-term projections and lower bounds on the size of the AIDS epidemic. Journal of the American Statistical Association, 83(402):301-308.

34. Levy, P. S. (1979). Samll area estimation-Synthetic and other procedures, 1968-1978. In J. Steinberg (Ed.), Synthetic Estimates for Small Areas. Nida Research Monograph 24. Department of HEW, Washington, D.C.: U.S. Government Printing Office.

35. Hamill, D. (1988). Indicator Approach. Paper presented at the Prevalence Estimation Technique conference held in Washington, D.C. November.

36. National Center for Health Statistics (1968). Synthetic State Estimates of Disability. PHS Publication No. 1759. Public Health Service, Washington, D.C.: U.S. Government Printing Office.

37. Demaree, R. G., and B. W. Fletcher (1981). Estimates of the Nationwide Prevalence of Heroin Use in 1981 and 1982. National Institute on Drug Abuse, Division of Epidemiology and Statistical Analysis.

38. Chandra-Sekar, and W. Edwards Deming (1949). On a method of estimating birth and death rates and the extent of registration. Journal of the American Statistical Association, 44:101-115.

39. Ericksen, Eugene P., and Joseph B. Kadane (1985). Estimating the population in a census year: 1980 and beyond. Journal of the American Statistical Association, 80(389):98-109.

40. Bishop, Y. M. M., S. E. Fienberg, and P. W. Holland (1975). Discrete Multivariate Analysis: Theory and Practice. Cambridge, Mass.: MIT Press.

41. Fienberg, S. E. (1972). The multiple recapture census for closed populations and incomplete 2^k contingency tables. Biometrika, 59:591-603.

42. Wickens, Thomas D. (1989). Multiway Contingency Tables Analysis for the social Sciences. Hillsdale, NJ: Lawrence Erlbaum Association.

43. Greenwood, J. A. (1971). Estimating the number of narcotic addicts. Drug Control division, Office of Scientific Support, Bureau of Narcotics and Dangerous Drugs, U.S. Department of Justice Document No. SCID-TR-3.

44. French, J. F. (1977). Prevalence of heroin use in New Jersey. Unpublished report of Department of Health, Alcohol, Narcotic, and Drug Abuse Unit, State of New Jersey.

45. Doscher, M. L., and J. A. Woodward (1983). Estimating the size of subpopulations of heroin users: Applications of log-linear models to capture-recapture sampling. The International Journal of the Addictions, 18:167-182.

46. Woodward, J. A., D. G. Bonett, and M. L. Brecht (1985). Estimating the size of a heroin abusing population using multiple-recapture census. In B. Rouse, N. Kozel, and L. Richards (Eds.), Self-report Methods of Estimating Drug Use: Meeting Current Challenges to Validity, pp. 158-171. National Institute on Drug Abuse Research Monograph 57, DHHS Publication No. (ADM)85-1402. Washington, D.C.: U.S. Government Printing Office.

47. Blumenthal, Saul, Ram D. Dahiya, and Alan J. Gross (1978). Estimating the complete sample size from an incomplete poisson sample. Journal of the American Statistical Association, 73(361):182-187.

48. Greene, M. A., and S. Stollmack (1981). Estimating the number of criminals. In J. A. Fox (Ed.), Models in Quantitative Criminology, pp.1-24. N.Y.: Academic Press.

49. Woodward, J. Arthur, Lynn Brecht, and Douglas G. Bonett (1987). Statistical Analysis of Drug Abuse Indicators. Final Report to the Drug Enforcement Administration and the National Institute on Drug Abuse, NIDA No. RA-ND-86-2.

50. Research Triangle Institute (1988). Presentation at the Prevalence Estimation Technique Conference held in Washington, D.C., November.

51. Seber, G. A. F. (1982). The Estimation of Animal Abundance and Related Parameters. MacMillan.

52. Seber, G. A. F. (1986). A review of estimating animal abundance. Biometrics, 42:267-292.

53. Jolly, G. M. (1965). Explicit estimates from capture-recapture data with both death and immigration-Stochastic model. Biometrika, 52:225-247.

54. Jolly, G. M. (1982). Mark-recapture models with parameters constant in time. Biometrics, 38:301-321.

55. Seber, G. A. F. (1965). A note on the multiple recapture census. Biometrika, 52:249-259.

56. Cormack, R. M. (1979). Models for capture-recapture. In R. M. Cormack, G. P. Patil, and D. S. Robson (Eds.), Sampling Biological Populations. Statistical Ecology Series, Vol. 5. Fairland, MD: International Co-operative Publishing House.

57. Cormack, R. M. (1981). Loglinear models for capture-recapture experiment on open populations. In R. W. Hiorns and D. Cooke (Eds.), The Mathematical Theory of the Dynamics of Biological Populations, II. London: Academic Press.

58. Cormack, R. M. (1985). Examples of the use of GLIM to analyze capture-recapture studies. In B. J. T. Morgan and P. M. North (Eds.), Statistics in Ornithology. Lecture notes in statistics, #29. Berlin-Heidelberg: Springer-Verlag.

59. Wickens, Thomas D. (1990). Multiple-capture Estimation of Population Size with Dynamic State Change. Manuscript in preparation.

60. Wickens, Thomas D. (1982). Models for Behaviors: Stochastic Processes in Psychology. San Francisco: Freeman.

61. Huber, J. J. (1962). Trap response of confined cottontail populations. Journal of Wildlife Manage, 26(2):177-185.

62. Manly, B. F. J. (1971). Estimates of a marking effect with capture-recapture sampling. Journal of Applied Ecology, 8:181-189.

63. Wittes, J. T. (1974). Application of a multinomial capture-recapture model to epidemiological data. Journal of the American Statistical Association, 69(345):93-97.

64. Carothers, A.D. (1973). Capture-recapture methods applied to a population with known parameters. Journal of Animal Ecology, 42:125-146.

65. Gilbert, R. O. (1973). Approximations of the bias in the Jolly-Seber capture-recapture model. Biometrics, 29:501-526.

66. Levin, G., E. B. Roberts, and G. B. Hirsch (1975). The Persistent Poppy, 229 pp. Ballinger, Cambridge, Mass.

67. Gardiner, L. K., and R. C. Shrechengost (1985). Estimating heroin imports into the united states. In B. A. Rouse, N. J. Kozel, and L. G. Richards (Eds.), Self-Report Methods of Estimating Drug Use: Current Challenges to Validity. NIDA Research Monograph 57, DHHS Publication No.(ADM)85-1402. Washington, D.C.: U.S. Government Printing Office.

68. Gardiner, L. K., and R. C. Shrechengost (1987). A system dynamics model for estimating heroin imports into the United States. System Dynamics Review, 3(1).

69. Shreckengost, R. C. (1984). How many heroin users are there? Unpublished manuscript dated June 4, 1984.

70. Homer, Jack (1990). A System Dynamics Simulation Model of Cocaine Prevalence. Final Report to National Institute of Justice under Grant 87-IJ-CX-0042.

71. Forrester, J. W., and P. M. Senge (1980). Tests for building confidence in system dynamics models. TIMS Studies in Management Sciences, 14:209-228.

Table 1

Characteristics of Major Drug Indicators

Survey/Indicator System	Sample	Data Collection	Sampling	Geographical Coverage	Time Frame	Relevance to Drug Abuse	Comments
General Surveys							
1. National Household Survey (NHS)	5000/8000 U.S. residients 12 years and older	Self-report	Stratified multistage	National	Cross-section every 2/3 year	Use of alcohol, tobacco, licit & illicit drugs in last 30 days, last year, lifetime	Refusal rates 16-23%; no coverage of high risk groups
2. Epidemiological Catchment (ECA)	from households & institutions, 18 years & older	Self-report	Multistage probability	Specific areas in St. Louis, Baltimore, New Haven, Raleigh-Durham, Los Angeles	Longitudinal– 3 interviews at 6-month intervals in 1980-84	Prevalence of mental disorders, including drug problems	Refusal rates 20-23%; data more representative of local areas & institutional subgroups than NHS
3. High School Senior Survey or Monitoring the Future	17000 high school seniors / young adults, college students	Self-report	Multistage-representative of all high schools in U.S.	National	Annual since 1975	Drug-related attitudes, opinions, use of various drugs	No coverage of dropouts, thus conservative estimates. Reliable for assessing trends
4. National Ambulatory Medical Care Survey	71,000 patient records from 2900 physician offices	Report by physicians	Three -stage Stratified Cluster	National	Annually for 1973-1982 and 1989	Diagnosis & treatment related to drug & alcohol abuse	Biased if high risk groups do not typically use physician offices; useful in assessing drug-related morbidity
5. National Hospital Discharge Survey	7,014,000 discharge records from hospitals	Report by hospitals	Two-stage Stratified	National	1963-1986	Prevalence of drug-related diagnoses in hospital patients	Biased if high risk groups do not typically use hospitals; useful in assessing drug-related morbidity
National Indicators							
6. Drug Abuse Warning Network (DAWN)	Case reports of emergency room(ER) visits & medical examiner (ME) mentions from 24-27 SMSAs	Report by emergency rooms and medical examiners	All non-federal short stay general hospitals with emergency rooms open 24 hrs/day, & all medical examiners in 24-27 SMSAs	Major SMSAs nationwide	Continuous	ER visits and ME mentions where drug abuse involved	Non-coverage of federal hospitals, children under 6, and alcohol-only incidence; not person-based

Table 1. (Continued)

Survey/Indicator System	Sample	Data Collection	Sampling	Geographical Coverage	Time Frame	Relevance to Drug Abuse	Comments
7. Vital Statistics	All drug-related deaths	Reports to Centers for Disease Control	All cases	National	Continuous	Drug-related deaths, AIDS, hepititis-B	Drug use information is not available unless is directly related to death
8. Uniform Crime Reports (UCR)	Summary of arrests (incidence of crime, most serious reported)	Aggregate by agency	Voluntary	National	Continuous	Drug-related crimes	Only most serious crime reported in multiple charge cases (drug crime are usually less serious); Aggregate summaries only--not case by case; incidence, not person-based
9. Drug Use Forecasting (DUF)	225 male and 100 female arrestees per study site	Self-report by interview & objective testing	The order of sampling priority is: non-drug and then drug offenses	Selected, nationwide--more than 20 largest cities	Quarterly since 1984--recently increased coverage	Drug use	Not probability sample--priority to non-drug arrestees; Objective measure (urine test) is a validity check on self-report; Detailed data which can be mapped onto UCR
10. Client-oriented Data Acquisition Process (CODAP)	Admissions to federally-funded drug treatment	Admission/ discharge records	All cases/currently voluntary reporting by treatment program	Selected nationwide	Continuous since 1972/major changes in reporting base in 1982	Primary, secondary, & tertiary drug use reason for admission to treatment	Limit by number of available treatment slots; Major changes in reporting base in 1982; No coverage of private treatment programs; Incidence based--no easy way to determine number of individuals responsible for admissions
11. STRIDE	Price/purity	Seizure/buy	By DEA offices	Selected Nationwide	Continuous	Price/purity of the drug	No distinction between wholesale and retail price

Table 2. Prevalence Estimation Techniques:
Utilities, Strengths and Difficulties

Method/Utilities	Strengths	Difficulties
Static Models		
Synthetic Estimation Making estimates in unknown areas by extrapolating from areas where prevalence is available or is known by another independent method. The population projection method extrapolates by mapping prevalence rates onto demographic characteristics of the target population. The principal component method extrapolates using a drug problem index that is derived by principal component analysis of several indicators.	-Requires little knowledge about the process -Relatively free from structural models -Requires fewer data sources of indicator data	-No structural properties -Estimates only as good as calibration sample -Make strong assumption of linearity and appropriate measures
Closed-population multiple capture Based on probability sampling theory, when two or more methods have been used to sample the relevant population during the same time frame, the relative sizes of the samples and their overlaps allow estimates to be made of the number of individuals that have not been detected.	-Statistically based -Integrates data from different survey methods -Provides statistical errors of estimation	-Require specific form/type of data (e.g., matching individuals across data sources) -Dependence on simplified probability model of independent and identical observations

Table 2. (Continued)

Method/Utilities	Strengths	Difficulties
Dynamic Models		
Open population multiple capture Based on probability sampling theories, the multiple-capture history observed over time (e.g., in treatment admission indicator) is used to generate estimates of the size of a partly hidden population.	-Focuses on minimal dynamic structure -Statistically based -Can describe changes in prevalence over time -Provides statistical errors of estimation	-Require specific form/type of data (e.g., matching individuals across time) -Dependence on simplified probability model (e.g., identical and independent sampling probability, etc.) -Cannot estimate sizes of unsampled (hidden) populations
System Dynamics Establishes a system connecting all relevant sources of data over time using feedback loops that are responsible for observed systematic changes. These relationships provide estimates of missing observations and can be projected to the near future.	-Provides comprehensive description of the processes -Has a general dynamic structure -Can estimate sizes of incompletely observed populations	-Difficult to build, calibrate and validate -May contain structures not supported by data -Generalizability of model is likely restricted