U.S. Department of Justice
Office of Justice Programs
National Institute of Justice

NATIONAL INSTITUTE OF JUSTICE

Research Report

Techniques for the Estimation of Illicit Drug-Use Prevalence:

Overview of Relevant Issues

About the National Institute of Justice

The National Institute of Justice is the research and development agency of the U.S. Department of Justice, established to prevent and reduce crime and to improve the criminal justice system. Specific mandates established by Congress in the Omnibus Crime Control and Safe Streets Act of 1968, as amended, and the Anti-Drug Abuse Act of 1988 direct the National Institute of Justice to:

- *Sponsor special projects and research and development programs* that will improve and strengthen the criminal justice system and reduce or prevent crime.
- *Conduct national demonstration projects* that employ innovative or promising approaches for improving criminal justice.
- *Develop new technologies* to fight crime and improve criminal justice.
- *Evaluate the effectiveness of criminal justice programs* and identify programs that promise to be successful if continued or repeated.
- *Recommend actions* that can be taken by Federal, State, and local governments as well as private organizations to improve criminal justice.
- *Carry out research on criminal behavior.*
- *Develop new methods of crime prevention* and reduction of crime and delinquency.

The National Institute of Justice has a long history of accomplishments, including the following:

- Basic research on career criminals that led to development of special police and prosecutor units to deal with repeat offenders.
- Research that confirmed the link between drugs and crime.
- The research and development program that resulted in the creation of police body armor that has meant the difference between life and death to hundreds of police officers.
- Pioneering scientific advances such as the research and development of DNA analysis to positively identify suspects and eliminate the innocent from suspicion.
- The evaluation of innovative justice programs to determine what works, including drug enforcement, community policing, community anti-drug initiatives, prosecution of complex drug cases, drug testing throughout the criminal justice system, and user accountability programs.
- Creation of a corrections information-sharing system that enables State and local officials to exchange more efficient and cost-effective concepts and techniques for planning, financing, and constructing new prisons and jails.
- Operation of the world's largest criminal justice information clearinghouse, a resource used by State and local officials across the Nation and by criminal justice agencies in foreign countries.

The Institute Director, who is appointed by the President and confirmed by the Senate, establishes the Institute's objectives, guided by the priorities of the Department of Justice and the needs of the criminal justice field. The Institute actively solicits the views of criminal justice professionals to identify their most critical problems. Dedicated to the priorities of Federal, State, and local criminal justice agencies, research and development at the National Institute of Justice continues to search for answers to what works and why in the Nation's war on drugs and crime.

Techniques for the Estimation of Illicit Drug-Use Prevalence:

An Overview of Relevant Issues

May 1992

NCJ 133786

Yih-ing Hser, M. Douglas Anglin, Thomas D. Wickens,
Mary-Lynn Brecht, and Jack Homer

National Institute of Justice

Charles B. DeWitt
Director

Michael J. Russell
Deputy Director

Prepared for the National Institute of Justice,
U.S. Department of Justice, by the UCLA Drug Abuse Research Group,
under contract number 90–IJ–CX–0014.

M. Douglas Anglin, Ph.D., is Adjunct Associate Professor with UCLA's Neuropsychiatric Institute and an internationally known drug abuse researcher. Mary-Lynn Brecht, Ph.D., Senior Statistician at UCLA, has been working with multiple recapture models. Jack Homer, Ph.D., is an expert in system dynamics. Yih-Ing Hser, Ph.D., is Adjunct Assistant Professor at UCLA's Neuropsychiatric Institute and is specializing in quantitative psychology. Thomas D. Wickens, Ph.D., is Professor of Psychology at UCLA. His specialty is quantitative psychology with emphasis on the Markov and multiple recapture models.

Foreword

Accurate and timely knowledge of the size and characteristics of drug-user populations is vitally important for informed public policies aimed at the control of drug abuse. It supports the monitoring of trends in drug-related problems, the development of options for prevention and treatment, the projection of criminal justice needs, and the evaluation of policy effectiveness.

The development of the tools for amassing this information has long been of mutual interest to the National Institute of Justice (NIJ) and the National Institute on Drug Abuse. NIJ has funded a series of studies on the development of statistical models for prevalence estimation.

This report results from an NIJ study of cocaine prevalence by the University of California, Los Angeles, Drug Abuse Research Group. It surveys the complexities of defining drug-use prevalence, of getting reliable new data, and of creating trustworthy models to generate prevalence estimates from the data. Our report clearly shows that practical estimates of crime and drug abuse require careful development and thorough testing.

At NIJ, our goal is to provide a scientifically defensible basis for these estimates, so that they can be confidently used for policy planning and resource allocations.

Charles B. DeWitt
Director
National Institute of Justice

Acknowledgments

The authors are members of the UCLA Drug Abuse Research Group and constitute a team working on two successive projects sponsored by the National Institute of Justice: Cocaine Prevalence Estimation (87–IJ–CX–0042) and Policy Evaluation (90–IJ–CX–0014). The project is led by Dr. Anglin as Principal Investigator, Dr. Hser as Project Director, and Drs. Wickens, Brecht, and Homer as experts for the individual methodologies under investigation. Project products are a team effort. Dr. Hser and Dr. Anglin are also supported in part by Research Scientist Development Awards (DA00139 and DA00146, respectively) from the National Institute on Drug Abuse.

Further support was obtained from contract D–0001–09 from the California Department of Alcohol and Drug Programs. We wish to recognize several government agencies for their generous support of the present research and the individual researchers whose insightful comments and suggestions contributed to the work of the team. In particular, we wish to thank Dr. Bernard Gropper, Director of the Drugs and Crime Program at the National Institute of Justice, for his full support of our prevalence estimation research, from which we have developed the present review. The National Institute on Drug Abuse kindly supplied data tapes on the National Household Survey on Drug Abuse and the Drug Abuse Warning Network; Mr. Mark Brodsky was particularly helpful in this respect. The Federal Bureau of Investigation provided Uniform Crime Reports data, and the Drug Enforcement Administration provided data on the System To Retrieve Drug Evidence. Dr. Peter Reuter of the RAND Corporation reviewed previous versions of this manuscript. 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.

Table of Contents

Executive Summary

Estimation of the size of drug-using populations is important for policy decisions concerning crime control, public health, and the allocation of intervention resources. Such estimation is particularly difficult because drug use is both a stigmatized and an illegal behavior.

Most general surveys do not adequately access groups that are at high risk for drug use and often rely only on self-report. More objective data are typically available only for nonrandom samples of various populations. Accurate estimates can be obtained only by the careful application of appropriate methodologies that consider both valid descriptions and measures of the phenomena and that assess the size of hidden or unknown populations.

Methodological issues for obtaining reliable and valid estimates and their interpretation—including adequate definition of populations, the availability and suitability of existing data, and the utility and applicability of prevalence estimation techniques—are the primary focus of this review. The review is written for policymakers and other readers interested in evaluating the adequacy of estimates of illicit drug users. It is also intended to assist anyone interested in doing prevalence estimation in the choice of accessible data and appropriate methods.

Population Coverage and Classification Considerations

Prevalence can refer to people, events or occasions, the quantity of drugs consumed, or dollars used to finance the drug trade. In this review, we are concerned mostly with counting people. Prevalence estimation must also consider timeframe and geographic boundaries because drug use is a dynamic process: nonusers move into the actively using population while current users may cease use, and drug users sometimes move in and out of targeted geographic areas.

Types of drugs and levels of use often need to be considered for several reasons. First, multiple-drug use is common; therefore, estimates of users need to consider this overlap of person and multiple consumption occasions to avoid repeated counts of individuals. Second, because of the associated severe social and health consequences, some prevalence estimation may focus on a particular type of drug user (e.g., intravenous-drug users) or more severe levels of use (e.g., addicts). However, when prevalence estimation is approached from a view of total drug consumption, all categories of drug users must be considered because all levels of drug use contribute. Definitions are necessary to identify appropriate data sources and to provide valid prevalence estimates.

Data Conventionally Used for Prevalence Estimation

Several national surveys and special-purpose Federal data systems contain drug-related information and have been conventionally used for prevalence estimation. Major surveys include the National Household Survey on Drug Abuse, the High School Senior Survey, the National Ambulatory Medical Care Survey, and the National Hospital Discharge Survey.

Executive Summary (continued)

Examples of Federal data systems include the Drug Abuse Warning Network, the System To Retrieve Drug Evidence, the Uniform Crime Reports, the Drug Use Forecasting program, and the Client-Oriented Data Acquisition Process. These data sources have both strengths and weaknesses, and several issues must be considered in assessing the utility of a particular data system for prevalence estimation purposes.

Sampling

Most large-scale surveys of drug use are based on probability sampling techniques. However, general surveys often undersample certain high-risk groups because of nonresponse or noncoverage. If drug users are disproportionately overrepresented among the nonrespondents, prevalence rates determined from the data are too small. Similar biases apply if coverage of high-risk groups is inadequate. Other Federal indicator systems that do not use survey methodology are typically not probability samples and may manifest severe coverage limitations.

Validity of Drug-Use Indicators

For legal or social reasons, the disclosure of stigmatized behavior, such as illicit drug use, is generally resisted. Memory failure and other cognitive complexities in recalling behavior can also distort self-reported drug use. Validation of self-report in such general surveys is difficult and is generally lacking. Other types of indicators, such as urinalysis, can provide some limited validity check on self-reported recent drug use. However, the window of detectability in urine varies with the type of drug tested, and it is typically only for a few days.

Event-Based and Person-Based Data Systems

Event-based systems are those in which each record arises from a single event, such as 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 data, most Federal monitoring systems are event-based recording systems. Often, several records belong to one individual who has multiple treatment admissions or emergency room episodes. Thus, the number of people actually responsible for the number of records in these data needs to be determined if the person-based prevalence estimate is desired.

Data System Consistency

The description of trends in use over time requires consistency in the reporting panel—a goal not often achieved in continuously reporting indicator systems over long periods of time, especially when reporting by the contributing agencies is voluntary. Inconsistencies in reporting standards and practices must also be con-

sidered. Finally, the various data systems may not share standardized methods of data collection, so that the comparison of indicators and their interpretation must be carefully attended.

Prevalence Estimation Methods

We selected for indepth description and discussion several prevalence estimation methods for their historical importance and promising future applications. These methods include *synthetic estimation* (population projection models and principal components approaches), *multiple-capture census* (closed-population capture models, open-population capture models such as the truncated Poisson estimation model, ecological open-population models, and Markov-based dynamic recapture models), and *system dynamics* modeling.

A useful distinction among these methods is to categorize them as primarily static versus dynamic approaches. *Static methods* such as synthetic estimation or closed-population models evaluate a population at a single point in time. *Dynamic methods* such as open-population models or system dynamics models are appropriate when a population is traced over time. The choice of model is largely dictated by the data available.

Static Models

The least complex of the prevalence estimation models are those that describe the system at a single point in time. Because of their convenience, these models have been the most popular in drug-use prevalence estimation.

Synthetic estimation. The simplest of the static estimation techniques are those that employ synthetic estimation. These methods develop prevalence estimates for new populations using several more readily available data sources or indicators from known populations by matching various predictor variables, usually demographic characteristics (e.g., ethnicity, gender, age, and regional location) and determining appropriate weighting schemes. The following two methods are used to determine the proper weights to be applied.

Population projection models are based on the logic that if the drug-use prevalence rates are known in a population having known demographic distribution, then the relationships between prevalence and demographic characteristics can be transferred to another population, either smaller or larger than the first.

The *principal components approach* uses the relationships observed among multiple indicators in several geographic areas, such as Standard Metropolitan Statistical Areas, in an attempt to obtain a single composite and common indicator of drug use. By combining several indicators into a single composite index with the weights determined by the principal components analysis, an index that is more reliable than any single indicator alone may be derived.

Closed-population capture models. Another variety of static estimation procedure is *closed-population multiple-capture methods*. In these techniques, two or more surveys of different data sources such as emergency room and arrest records are used to probe the given population during a single timeframe. Each survey must be able to identify individual cases and determine which individuals have been detected in which of the surveys.

Using the information about the relative sizes of the samples and their overlaps, one can estimate the number of individuals who have not been detected. The population whose size is to be estimated here is said to be closed because of the single sampling time.

Dynamic Models

Dynamic models differ from the static models in that they describe processes over time. The process examined may be the states of individual drug users, as in the open-population multiple-capture models, or it may be a collection of aggregate societal states, as in the system dynamics models. Because these models can represent processes more accurately than static models and because they typically incorporate more information, these models potentially result in more accurate estimates of prevalence.

Open-population capture models. Data that are drawn from identifiable individuals in a system of successive surveys or censuses using the *open-population multiple-capture method* are similar to those obtained in the closed-population multiple-capture situation described earlier. However, open-population models keep track of population changes and provide estimates reflecting such time-related changes of in-and-out flows. The data from a longitudinal dynamic process thus require a different type of statistical model, based on some assumptions about the open population under consideration.

The *truncated Poisson estimation model* is the simplest version of the multiple-observation models. It can be applied in situations where only the frequency that an individual appears in a data system is recorded. The result is a frequency distribution for a given period of time, starting with the count of individuals observed once and continuing upward. The unobserved portion of the population resides in the missing "zero" cell of this distribution (i.e., those individuals never observed) and is estimated.

A more sophisticated approach utilizing repeated sampling may provide a better estimation methodology. *Ecological open-population models* assess the size and character of a biological population based on repeated marked samples. The most common class of ecological sampling models that are applicable to the estimation of the number of drug users is the capture-recapture type. However, some of the assumptions of these models are unrealistic for drug-using 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.* Instead of counting the individuals captured 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 in which users are drawn from a large population of nonusers. After this first observation, the balance of the process is governed by the dynamics of a state structure that 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, and the procedures generate estimates of the size of the population from which the observations are drawn.

System dynamics modeling. System dynamics is a general methodology for analyzing dynamic phenomena through the use of simulation models based on information-feedback control theory. A system dynamics model consists of an interconnected set of difference equations representing continuous-time movement and accumulations of people, materials, and information. After being assigned initial conditions consistent with historical data, the set of equations is used to generate output over time. If the model is a valid one, this output will closely mimic the true course of events, and the model may be used for prevalence estimation and making conditional forecasts.

Summary

Prevalence estimation methods differ in their data requirements and statistical properties and thus in their utility for providing certain types of prevalence estimates. Each of these methods has strengths and limitations, and none can provide estimates without knowledgeable and careful application. Some of these limitations are due to the necessary simplicity of the assumptions of the model, while others are due to the demands for specific data of a certain quality. Therefore, in selecting a model, the user must consider the appropriate use of the data available and interpret estimates derived from these methods within the appropriate context.

Concluding Comments

The quality of prevalence estimates is the result of an interplay among theory about the phenomena, the estimation methodology, and the empirical data. Uncertainties in any of these areas affect the accuracy of the results. Given the current level of knowledge about drug use and the available data, multiple methods using different approaches and data sources are necessary to provide estimation ranges that set boundaries for policy decisions. Continued and consistent efforts in improving the understanding of drug-use phenomena, the quality of data collection systems and prevalence estimation techniques, and their appropriate utilization are necessary to ensure more valid estimation results.

Introduction

Although illicit drug use is widely perceived as one of the Nation's most serious problems, only recently have several estimates of its extent attracted considerable public attention. For example, the Office of National Drug Control Policy (ONDCP) estimated in its 1990 report that there were 860,000 hardcore cocaine addicts in the United States (ONDCP 1990). Shortly afterwards, and in sharp contrast, the Senate Judiciary Committee (1990) announced their estimated number of hardcore cocaine addicts to be 2.2 million.

Some researchers (e.g., Wish 1990) also estimated that between 978,000 and 1.3 million arrestees would have tested positive for cocaine in 1988 in the 61 largest cities in the United States. Asking a slightly different question, the Institute of Medicine (Gerstein and Harwood 1990) estimated that approximately 5.5 million Americans needed drug treatment on a typical day in 1987–1988. The Office of Technology Assessment (1990) placed the estimated number of intravenous-drug users in the United States at between 1.1 and 1.8 million. The proliferation and divergence of these estimates for various types of drug users underline the considerable policy interests and diverse needs for obtaining accurate drug-use prevalence estimates.

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 the suitable targeting of prevention programs. However, methodologies for obtaining accurate estimates are difficult to operationalize and to implement. Because of the illegal nature of drug use, prevalence estimates based on self-report surveys suffer from considerable reliability and validity problems, while more objective data are typically available only for nonrandom samples of the population. Therefore, drug-use prevalence has often been defined on an ad hoc basis and has been measured in different ways. As a result, there is no clear basis by which to compare estimates.

Assessment of drug-use prevalence cannot be made without addressing several basic issues. The purpose of this paper is to answer critical questions by reviewing and explicating relevant issues: How has drug-use prevalence been defined, and in what ways are these definitions applied in practice? What are the characteristics of available data sources, and how do they support or limit prevalence estimation? What estimation techniques are available to support prevalence estimation, and for what contexts are they appropriate? The rest of the paper presents materials addressing issues of population coverage and classification, data sources, and those prevalence estimation techniques that appear most promising to pursue.

The focus of the paper is on methodology, not on the resulting estimates themselves. Only by using an appropriate methodology to assess a phenomenon can we obtain reliable and valid estimates. The critical issues raised here should be considered in making judgments about the adequacy of any estimation result. The discussion is intended, in part, to assist people interested in doing prevalence estimation in their choice of accessible data and appropriate methods.

Population Coverage and Classification Considerations

One primary issue for any prevalence estimation effort is simply what to measure. In the context of drug use, prevalence can refer to people, events or occasions, the quantity of drugs consumed, or dollars used to finance the drug trade. In the present paper, we are concerned mostly with counting people.

Specifications for prevalence estimation must also consider timeframe and geographic boundaries. These last two considerations are important because drug use is a dynamic process; nonusers move into the actively using population while current users may cease use—temporarily or permanently—due to treatment, incarceration, or death. Drug users also move into and out of geographic areas. Specifications for time intervals or geographic areas are usually determined for practical or jurisdictional reasons. The following section presents several of the most commonly used categories based on time and geographic consideration.

Temporal and Geographic Considerations

Determining inclusion within the population of interest based on past use or current use constitutes the minimal temporal criterion of prevalence. Several prevalence categories are classified by temporal criteria and 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 subpopulation of current users 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 1-year period.

Geographic areas are typically defined in a straightforward manner: a nation, a State, a region, a county, a city, or Standard Metropolitan Statistical Areas (SMSA's). Choices of geographic area are usually made on jurisdictional grounds or for areas where sufficient data are available to support prevalence estimation.

Of most concern, and perhaps the most difficult definitional task, is the specification of the nature of the drug-using population because of the variation in drug-use levels among users. Three major problems arise: (1) the nature of the drug and consequences of its use, especially the type or level of use that is of concern; (2) the definition of categories, such as "addicts," "occasional users," etc.; and (3) the practical problem of accessing the appropriate sample. The following section

Population Coverage and Classification Considerations (continued)

presents a brief review of the most commonly used categories and definitions of drug-use patterns and consequences.

Conceptions of Drug-Use Patterns and Consequences

Specification of the drug of interest is necessary to identify appropriate 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 (Wesson and Smith 1979).

The most popular example is the use of a drug in combination with alcohol. Use of marijuana in combination with use of "hard" drugs such as heroin, cocaine, or others is also common. 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 consider this overlap of person and multiple-consumption occasions to avoid repeated counts of individuals due to either ongoing or intermittent polydrug use.

For most drugs, patterns of use are described by frequency of use. For example, one schema characterizes patterns of use into five categories: experimental, recreational, circumstantial, intensified, and compulsive (Siegel 1984). Adolescents, for example, often engage in *experimental use* with psychoactive drugs of all kinds as they come in contact with specific drugs in their peer culture. *Recreational use* is the most common pattern of use; the main characteristic is self-control of consumption. Most marijuana and powder cocaine users fall into this category of use. People engaged in *circumstantial use* 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 immediate health effects or 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 and a higher probability of health consequences.

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, the availability and suitability of the data, and whether one particular schema is better suited for the estimation method 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 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 deescalation to lower levels—has important implications for developing appropriate models 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 1 month. Powder cocaine users rarely report this rapid escalation to daily or near-daily use. Crack cocaine users, however, typically have a more rapid escalation to high levels of use that more closely approximates that of heroin addicts (Khalsa, Anglin, and Paredes 1991).

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 (Edwards, Arif, and Hodgson 1981). 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-related information.

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

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 general data systems (excepting general sample surveys) have been constructed using inclusion criteria 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 directed toward counting persons. 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 that differentially interact with drug use. As a result, the generalizability of estimates based on such data sources may be suspect.

Data Conventionally Used for Prevalence Estimation

Because the availability and suitability of data are important considerations in choosing prevalence estimation approaches and because data profoundly affect 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 drug 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 (NHS) on Drug Abuse, which has been conducted every 2 or 3 years since 1971 among about 8,000 household residents aged 12 and older. Since 1990, NHS has been conducted annually. Another important survey data source, the annual High School Senior Survey (HSSS) (Johnston, O'Malley, and Bachman 1989), consists of self-administered questionnaires completed by approximately 17,000 high school seniors. The National Ambulatory Medical Care Survey (NAMCS) (Bryant 1988, Nelson 1988) gathers information about patient visits to non-Federal, office-based physicians in the Nation. The National Hospital Discharge Survey (NHDS) (Graham 1988) 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), which contains drug-related hospital emergency room visits and medical examiner or coroner mentions; the System To Retrieve Drug Evidence (STRIDE), which provides information on seizures of illicit drugs; the Uniform Crime Reports (UCR) data base on arrests and convictions for illegal drug possession or trafficking; the Drug Use Forecasting (DUF) program, 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), which contains hospital or clinic records on people seeking treatment for drug-related problems. These data systems are often called drug indicators because they have been used to indicate various aspects of drug use.

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.

Sampling

Most large-scale surveys of drug use (e.g., NHS, HSSS) are based on probability sampling techniques. Because of data acquisition considerations and economic reasons, such surveys are seldom based on simple random sampling procedures in which there is an equal probability of selecting any respondent and independent selection among respondents.

Table 1

Characteristics of Major Drug Indicators

Survey/Indicator System	Sample	Data Collection	Sampling	Geographical Coverage	Timeframe	Relevance to Drug Abuse	Comments
General Surveys							
1. National Household Survey (NHS)	5,000/8,000 U.S. household residents 12 years and older	Self-report	Stratified multistage	National	Cross-section every 2–3 years (1971–1989), now annually	Use of alcohol, tobacco, and licit and illicit drugs in past 30 days, past year, lifetime	Refusal rates 16–23 percent; no coverage of high-risk groups not in households
2. Epidemiological Catchment (ECA)	From households and institutions, 18 years and 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 1980–1984	Prevalence of mental disorders, including drug problems	Refusal rates 20–23 percent; data more representative of local areas and institutional subgroups than NHS
3. High School Senior Survey or Monitoring the Future	17,000 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 (NAMCS)	71,000 patient records from 2,900 physician offices	Report by physicians	Three-stage stratified cluster	National	Annually for 1973–1982 and 1989	Diagnosis and treatment related to drug and alcohol abuse	Biased if high-risk groups do not typically use physician offices; useful in assessing drug-related morbidity
5. National Hospital Discharge Survey (NHDS)	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 and medical examiner (ME) mentions from 24–27 SMSA's	Report by ER's and medical examiners	All non-Federal short-stay general hospitals with ER's open 24 hrs/day, and all ME's in 24–27 SMSA's	Major SMSA's nationwide	Continuous since 1972	ER visits and ME mentions where drug abuse involved	Noncoverage of Federal hospitals, children under age 6, and alcohol-only incidence; not person-based; largely clinical signs and self-report, not lab confirmed
7. Vital Statistics	All drug-related deaths	Reports to Centers for Disease Control	All cases	National	Continuous since 1966	Drug-related deaths, AIDS, hepatitis-B	Drug-use information not available unless directly related as causes or contributing factors to death
8. Uniform Crime Reports (UCR)	Summary of arrests (incidence of crime, most serious reported)	Aggregate by agency	Voluntary	National	Continuous since 1966	Drug-related crimes	Only most serious crime reported in multiple-charge cases (drug crimes are usually less serious); aggregate summaries only—not case by case; incidence-based, not person-based; new incident-based UCR is currently under development

Table 1

Characteristics of Major Drug Indicators (continued)

Survey/Indicator System	Sample	Data Collection	Sampling	Geographical Coverage	Timeframe	Relevance to Drug Abuse	Comments
9. Drug Use Forecasting (DUF)	225 male and 100 female arrestees per study site	Urine tests and self-report by personal interview	Nonrandom samples; sampling priority given to nondrug offenses over drug offenses	Selected more than 20 largest cities nationwide	Quarterly since 1986 (sites added each year to 1990)	Drug use; related behaviors, such as needle use	Not probability sample of arrestees or representative sample of jurisdictions; 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, and tertiary drug use reason for admission to treatment	Limited 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. System To Retrieve Drug Evidence (STRIDE)	Price/purity	Seizure/buy	By Drug Enforcement Agency offices	Selected nationwide	Continuous since 1973	Price/purity of the drug	No distinction between wholesale and retail price and purity

Data Conventionally Used for Prevalence Estimation (continued)

However, most of these surveys have chosen sampling designs that achieve representativeness of the target population at the highest level of aggregation (e.g., household, regional, national). Although these types of probability samples may not meet the strict assumptions of some prevalence estimation methods, they are commonly used for estimation purposes.

NHS, for example, employs a stratified, multistage sample design based on a predetermined sequence of selection criteria to achieve a representative sample of households nationwide. 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 who 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 underlying distribution of the study population.

However, general surveys often undersample certain high-risk groups because of nonresponse or noncoverage. Given the difficulties of fieldwork, response rates of about 80 percent achieved by these surveys are a significant accomplishment. Nonetheless, the 20-percent nonresponse 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 1 percent within the general population are typical. If drug users are disproportionately overrepresented among the nonrespondents, prevalence rates are substantially underestimated.

Inadequate coverage of high-risk groups is another potential bias. For example, NHS excludes those in group quarters (military installations, correctional institutions, college dormitories, and hospitals) and those who have no permanent residence (the homeless and residents in single-room occupancy hotels). Thus, it is likely that NHS 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 permanent household living arrangement; these groups are at highest risk for illicit drug abuse (Gandossy, Williams, Cohen, and Harwood 1980; Robins and Wish 1977; Wish 1990). As one example, an analysis of illicit drug use among arrestees indicated there are two to six times more regular cocaine users in the arrestee population alone than NHS indicated for the whole Nation (Wish 1990). 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) (Eaton and Kessler 1985), 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 homeless shelters.

Limitations of sampling that result in underestimation of drug consumption are also exemplified in HSSS. Persons who dropped out of school prior to their senior year or students who were absent on the day of the survey are not included in the sample. It is estimated that the dropout rate in the United States may average 15 to 20 percent of a birth cohort (Johnston et al. 1988); it may be even higher in some urban minority populations. It is also known that drug use among dropouts and those frequently absent from school is higher and more extensive than among their peers who continue in school (Kandel and Maloff 1983). Therefore, estimates of drug use from these data sources are conservative. However, if these factors do not change over survey years, the relative trends in drug use may be reliable.

Other Federal indicator systems are not probability samples and may manifest severe coverage limitations. For example, the DAWN system, which reports drug-related emergency room visits and medical examiner mentions, covers only non-Federal hospitals. The DUF program, maintained by the National Institute of Justice, monitors drug use by arrestees and places priority on sampling arrestees whose charge is for a non-drug-related offense.

Validity of Drug-Use Indicators

The disclosure of stigmatized behavior such as illicit drug use to professional surveyors is generally resisted for legal or social reasons (Rouse, Kozel, and Richards 1985). With assured confidentiality and anonymity of 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) (O'Malley et al. 1984).

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 HSSS, 18 percent of the white and 28 percent of the black students stated in 1985 that they may not have reported heroin use even if they had used it (Johnston et al. 1988). Within the arrestee population, only about half the number of arrestees with positive urinalyses self-reported recent drug use (Harrison 1990).

Besides deliberate underreporting, memory failure and other cognitive complexities in recalling behavior can distort self-reported drug use. Validation of self-report in such general surveys is difficult (Harrell 1985). Urinalysis provides a limited validity check on self-reported drug use. The detection time varies with the type of drug tested, but it is typically only a few days.

Data Conventionally Used for Prevalence Estimation (continued)

Event-Based and Person-Based Data Systems

The distinction between event-based and person-based records contained in the available data systems must be considered in prevalence estimation, since this distinction influences the meaning of the count reported. 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 CODAP, are event-based record systems. Often, several records belong to one individual who has multiple treatment admissions or emergency room episodes. Thus, 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 for accuracy.

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 CODAP treatment admission records, based on several demographic, treatment, and drug use characteristics (Woodward, Retka, and Ng 1984). This method, however, is 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). The problem remains, however, that all the difficulties associated with self-reporting— such as memory failure or under- or overreporting—can bias the resulting estimates.

Another set of data systems, only alluded to earlier, are those concerned with quantity measures, rather than person- or event-based measures. For example, both the Customs Service and the Drug Enforcement Administration are concerned with the quantity of drugs illegally imported into the United States. These and other agencies are also concerned with the quantity of dollars that support the illicit drug economy. While estimation of quantity measures is important, such analyses are not the focus of this paper. (See, however, *Worldwide Cocaine Situation: 1990* 1991.)

Data System Consistency

Trend analysis requires consistency in the reporting panel, a goal 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 CODAP, a nationwide data base 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 withdrew from the system. However, this indicator system on treatment admissions remains comprehensive and historically 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 reduce the usefulness of the UCR data maintained by the Federal Bureau of Investigation 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 from certain types of arrests to others. Treatment availability may also change depending on allocated resources. Therefore, interpretation of these dynamic indicators is usually not straightforward and needs careful qualification.

Because the various data systems do not share standardized methods of data collection, comparison of indicators and the interpretation of their meanings can be impaired. The choice of appropriate models or methods must include consideration of the above-mentioned data limitations. Having explicated basic data issues, we next review the models judged most useful in prevalence estimation.

Whenever the size of a population must be estimated instead of being directly observed, it is necessary to make various assumptions about the phenomenon of interest, the population under investigation, and the observation procedure. Under most circumstances, these assumptions constitute either a mathematical or a statistical model. Using such models, incomplete information about the population is extrapolated to result in an estimate for the total population, or estimation is made for a new time period or geographic area for which such information is not directly available.

Several techniques have been applied to the problem of estimating the prevalence of drug use. The discussion below reviews both historical and promising prevalence estimation models of varying complexity. These methods are divided broadly into the *static* methods, which evaluate a population at a single point in time, and the *dynamic* methods, which trace a population over time.

Within these broad categories, we review in detail several specific methods, including synthetic estimation, multiple-capture census, and system dynamics modeling. Several models that are less frequently used in drug-use prevalence estimation, such as social network analysis (Frank 1979) and backwards extrapolation (Brookmeyer and Gale 1988), are not included in this discussion.

To avoid distraction by technical details, the relevant mathematical bases for the most important methods are presented in separate boxes. Readers interested in the technical aspects of the models should refer to the corresponding exhibit. Information on the more complicated models (e.g., the ecological open-population model) or conventional statistical models (e.g., principal components analysis) can be found in referenced articles or standard statistical textbooks.

The choice of model type is dictated largely by the data being used. When the data consist of several samples collected at about the same time, then a static model is appropriate. For data extending over several time periods, a dynamic model is appropriate. The line between the methods is not always clear. For example, a series of samples taken at nearly contiguous points in time is often analyzed by the static closed-population model without much error.

In general, the static models are 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. In contrast, the dynamic models have the potential to produce better estimates when they can be appropriately applied, but they require many more assumptions about the temporal evolution of drug-use patterns. To the extent that these assumptions are inaccurate, estimates of dynamic models may fail to a degree much greater than their statistical standard errors imply.

Prevalence Estimation Methods (continued)

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 tradeoffs between parsimony and validity. Thus, it is important to consider the underlying assumptions of the estimation approaches.

An important difference among the models is in their ability to identify so-called hidden populations, that is, populations of users who are severely undersampled or completely missing from the available data. Clearly, the ability to estimate the sizes of these hidden groups is critical to the prevalence estimation enterprise, but it is equally obvious that people or events never observed cannot be estimated without making some strong assumptions about the underlying process.

In general, the simpler models are less capable of adequately estimating these hidden populations, whereas the more complex and structured models are more likely to do a better job. However, it is also true that the more complex a model becomes, the more likely that some aspect of it incorrectly describes the true character of the population. Such model misspecification can undermine the validity of the estimates produced by the model. The tradeoff between parsimony and validity again applies.

Static Models

The least complex of the prevalence estimation models are those that describe the system at a single point in time. Because of their convenience, these models have been the most popular in drug-use prevalence estimation.

Synthetic Estimation

The simplest of the static estimation techniques are those that employ synthetic estimation. These methods develop prevalence estimates for new populations using several more readily available data sources or indicators from known populations by matching various predictor variables, usually demographic characteristics (e.g., ethnicity, gender, age, and regional location). The crux of the synthetic estimation methods is the selection of an appropriate set of predictor variables and the determination of the proper weights to be applied to them. The many alternative methods of weight determination that have been used range from simply transferring relationships found in one population to another population (Levy 1979) to some rationalized linear function (Hamill 1988) to factor analytic modeling (Person, Retka, and Woodward 1976; 1977).

Population projection models were originally developed by the National Center for Health Statistics (1968) for obtaining estimates of prevalence in easily defined areas such as cities or SMSA's (Levy 1979). The logic of these models is that if drug-use prevalence rates are known in one population having known demo-

Exhibit 1

Population Projection Models

(Synthetic estimation)

This procedure measures the number of individuals in a population that have some characteristic, such as drug use. The estimate is made by transferring information from a *calibration sample*, in which information about the prevalence of the characteristic has been measured, to a *target population*, where this information is not directly available. Linking demographic information—such as age, sex, or income level—must be available for both groups. For each combination of demographic variables **x**, one estimates the probability $P_c(\mathbf{x})$ that an individual in the calibration sample with this combination of attributes has the characteristic of interest. To make a synthetic estimate, these probabilities are assumed to hold in the target population for which only the frequencies $N(\mathbf{x})$ (or the proportions) of the demographic categories are known. For a given demographic cell, the number of members with the characteristic in the target population is estimated by the product of the number of members of the target population in this cell and the probability that a member has the characteristic. Using the calibration sample to estimate this probability, the product is $N(\mathbf{x})\,P_c(\mathbf{x})$. The estimate of the prevalence $\hat{N}$ in the new population is the sum of these products over the demographic categories,

$$\hat{N} = \sum_{\mathbf{x}} N(\mathrm{x})P_c(\mathrm{x})$$

The crucial assumption underlying this synthetic estimation procedure is that the probabilities $\mathbf{P_c(x)}$ determined in the calibration sample also apply to the new population.

graphic distribution, then the relationships between prevalence and demographic characteristics can be transferred to another population, either smaller or larger than the first (**see exhibit 1**).

Specifically, suppose 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 calibration sample, the proportion of drug-using members within each category is estimated. In a new population, a demographic survey indicates the frequencies in each combination of age, sex, and race. Combining this frequency information with the rates from the calibration sample (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 (Levy 1979).

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 NHS 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 obtained from NHS. More recent efforts (Wish 1990) have attempted to correct the deficiencies in NHS by combining its data with data derived from nonhousehold populations (e.g., the homeless or arrestees sampled by the DUF program). The basic methodologies applied, however, still use synthetic estimation procedures.

Despite its simplicity and widespread use, the use of the population projection method has been questioned for several reasons. A potential problem lies in the quality of the calibration sample data, which typically relies on survey-based prevalence data. Survey data based on the self-report of stigmatized behavior are

Prevalence Estimation Methods (continued)

particularly subject to sampling bias and distorted reporting. Another possible problem is that other unmeasured characteristics may make the rates in the target population different from those in the calibration sample.

For example, the method will not produce valid estimates when regional differences in drug use and availability render the demographic variables inadequate to yield an appropriate estimate. Regional differences are considerable in many indicator data systems. The direct form of synthetic estimation should be applied only when the calibration samples are sufficiently representative so that the observed pattern can be convincingly projected to a population of interest.

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., SMSA's) in an attempt to obtain a single composite and common 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 the underlying construct of the prevalence of drug use in the area. By combining these indicators with appropriate weighting into a single composite index, one can, in principle, derive an index that is more reliable than any single indicator alone.

The "Heroin Prevalence Index" (HPI) of Person, Retka, and Woodward (1976, 1977) illustrates the principal components procedure based on rank orderings of SMSA's 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 the other contributing or new geographic areas (e.g., other SMSA's).

The usefulness of the HPI approach depends largely on the acceptance of its basic assumptions. It is worth examining these assumptions in some detail, since the method's underlying problems are similar to those noted elsewhere for other estimation methods.

Difficulties can arise at each of the three steps. In the first step, the principal components approach assumes the measures are monotonically related. Except for measurement error, the rank ordering of the sampling units (e.g., SMSA's) 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 and Fletcher (1981). For example, given limited treatment resources, the probability of admission for any one heroin user may decrease due to the large number of heroin users there are, or simply because the reporting bases underlying the indicators are different among SMSA's. Moreover, unless

the constituent variables are standardized, a principal components 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 resource availability of treatment; where treatment access is limited, treatment admissions cannot rise in relation to the need for treatment.

The second step requires the HPI to be calibrated to match independent estimates of prevalence in two or more SMSA's. This matching can be no better than the quality of these independent estimates. Unfortunately, well-based independent estimates are not usually 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, SMSA's. Finally, use of the HPI assumes a linear relationship between the indicators and prevalence. The straight-line nature of this linking function 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 the availability of high-quality independent estimates in two or more equivalent geographic areas for reference points.

Closed-Population Capture Models

Another variety of static estimation procedure is *closed-population multiple-capture methods*. In these techniques, two or more surveys of different data sources such as emergency room and arrest records are used to probe the given population during a single timeframe. Each survey must be able to identify individual cases and determine which individuals have been detected in both of the 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" (Chandra-Sekar and Deming 1949; Ericksen and Kadane 1985). The cross-sectional nature of these closed-population models contrasts with the longitudinal open-population models discussed later.

Exhibit 2

Closed-Population Capture Models

(The Petersen estimate)

The simplest closed-population multiple-capture estimate applies when a single population (e.g., drug users) of unknown size N is observed in two different ways, creating two samples of its members. Suppose that n_1 individuals are observed in the first sample and n_2 are observed in the second sample. Of these individuals, n_{12} appear in both samples. The probability that an individual appears in the second sample is estimated by

$$P_2 \doteq \frac{n_2}{N}$$

Since N is unknown, this proportion cannot be calculated directly. However, the data do allow one to find the proportion of the first sample that reappears in the second sample,

$$p_2 = \frac{n_{12}}{n_1}$$

If the population is homogeneous and the two sampling operations are independent of each other, then p_2 is an estimate of P_2. Equating P_2 and p_2 and solving for N provides an estimate of the original population size,

$$\hat{N} = \frac{n_1 n_2}{n_{12}}$$

This estimate was originally developed by Petersen (1894). The crucial assumptions underlying the Petersen estimate are the homogeneity of the population and the independence of the two samples.

Prevalence Estimation Methods (continued)

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 two-by-two 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. The relationships among the samples and resampled members allow an estimate of the total population size to be modeled **(see exhibit 2).**

With more than two surveys, a higher dimensional table is obtained, always with one missing cell for those individuals not detected in any survey. To obtain an estimate of the total prevalence, the frequency in the unobserved cell is extrapolated from a log-linear model fitted to the observed cells (Bishop, Fienberg, and Holland 1975; Fienberg 1972; Wickens 1989) **(see exhibit 3)**. The estimation of the unobserved frequency usually is accomplished by fitting a simple association model to the incomplete table. The more samples that are available, the more complex this model can be. This type of model has been used by the Census Bureau for population estimation of individuals missed by the census (Ericksen et al. 1985).

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 a way to extrapolate to poorly measured populations. They are limited, however, by their rather restrictive data requirements. They require a series of surveys or observations, and each member

Exhibit 3

Incomplete-Table Estimation in a Closed Population

The idea of several censuses or observations of a closed population that is used for the Petersen estimate (exhibit 2) can be extended to accommodate more than two samples. Several samples are taken, all drawn from the same population, and the individuals are classified by the samples in which they appear. The number of individuals with each capture history is counted. For example, suppose that three samples are used and let n_{ijk} denote the frequencies of the particular sampling patterns; thus, n_{101} is the number of individuals that appeared in the first and third sample but not in the second. The population size N is the sum of all the n_{ijk}. However, this sum cannot be calculated because n_{000}, which is the number of individuals who were never observed, is unknown. The frequencies n_{ijk} form a three-way contingency table with one missing cell,

Second sample		Yes		No	
Third sample		Yes	No	Yes	No
First sample	Yes	n_{111}	n_{110}	n_{101}	n_{100}
	No	n_{011}	n_{010}	n_{001}	n_{000}

An estimate of n_{000} is made by fitting the observed portion of the table by a log-linear model, then extrapolating this model to the missing cell. Estimation methods of this type were proposed by Fienberg (1972); see Bishop, Fienberg, and Holland (1975) or Wickens (1989) for a discussion of the fitting procedures. The crucial assumption that underlies this method is the adequacy of the model. In particular, the unobserved members of the population must have sampling characteristics similar to the observed members. There cannot be a hidden portion of the population, i.e., one that is never at risk of being sampled by any of the observation methods.

Models of various complexity can be fitted to the table. The simplest model is one that treats the samples as mutually independent, as described by the log-linear model

$$\log \mu_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C$$

In a two-sample table, this model is equivalent to the assumptions of the Petersen estimate and gives the same estimate $\hat{N}$. When more than two samples are available, models that allow associations among the samples can be fitted. For example, all pairwise associations among three samples are included in the model

$$\log \mu_{ijk} = \lambda + \lambda_i^A + \lambda_j^B + \lambda_k^C + \lambda_{ij}^{AB} + \lambda_{ik}^{AC} + \lambda_{jk}^{BC}$$

Presumably, better models give more accurate estimates. To estimate the missing cell, the estimated parameters $\hat{\lambda}$, $\hat{\lambda}_i^A$, etc. that are associated with nonobservation are substituted in the relevant model. For example, using the four-parameter mutual-independence model,

$$\hat{n}_{000} = \exp[\hat{\lambda} + \hat{\lambda}_0^A + \hat{\lambda}_0^B + \hat{\lambda}_0^C]$$

The population size is estimated by the sum of $\hat{n}_{000}$ and the number of individuals actually observed.

Whenever the model has fewer parameters than there are distinct frequencies in the data table, its adequacy can be tested using a goodness-of-fit statistic that compares the observed n_{ijk} to the mean frequencies $\hat{\mu}_{ijk}$ estimated from the fitted model. Common measures of this type are the Pearson statistic,

$$X^2 = \sum_{\text{cells}} \frac{(n_{ijk} - \hat{\mu}_{ijk})^2}{\hat{\mu}_{ijk}}$$

and the likelihood-ratio statistic,

$$G^2 = 2 \sum_{\text{cells}} n_{ijk} \log \frac{n_{ijk}}{\hat{\mu}_{ijk}}$$

The sums are over the cells of the incomplete table. These statistics are referred to as a chi-square distribution with degrees of freedom equal to the difference between the number of frequencies in the incomplete table and the number of parameters in the model. For example, with three samples, seven frequencies are observed. The mutual-independence model has four parameters and can be tested with three degrees of freedom. The pairwise-association model has seven parameters and can be fitted, but the accuracy of the fit cannot be tested.

of the population must be equivalently at risk for detection in each sampling. The static models may be inappropriate when the samples are obtained at different times, especially if the intersample time period is lengthy.

Other difficulties in applying closed-population multiple-capture models in general include lack of comprehensiveness in coverage by all surveys, problems in matching individuals across samples, and lack of independence in the observations. 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 (Bishop et al. 1975).

Dynamic Models

Dynamic models differ from the static models in that they describe processes over time. The process examined may be the states of individual drug users, as in the open-population multiple-capture models, or it may be a collection of aggregate societal states, as in the system dynamics models. Because these models can represent processes more accurately than static models and because they typically incorporate more information, they have the potential to result in a more accurate estimate of prevalence.

Open-Population Capture Models

Data drawn from identifiable individuals in a system of successive surveys or censuses using the *open-population multiple-capture method* are similar to those obtained in the closed-population multiple-capture situation described earlier. These data also form an incomplete contingency table. The similarity is superficial, however, because 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 repeated censuses, some individuals leave the population before the final census and 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 of in-and-out flows. The data from a longitudinal dynamic process thus require a different type of statistical model, based on some assumptions about the open population under consideration.

Most applications of the multiple-capture methodology to longitudinally repeated drug-use samples have used closed-population models (e.g., Doscher and Woodward 1983; French 1977; Greenwood 1971; Woodward, Bonett, and Brecht 1985). Such estimates are potentially biased because some individuals may not be

available throughout the entire sequence of time sampling and because 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. Such poor estimation is especially serious within geographic areas representing relatively small populations. Models that accommodate the more realistic open-population dynamics may be generally more appropriate for estimation purposes.

The *truncated Poisson estimation model* is the simplest version of the multiple-observation models. It can be applied in situations where only the frequency that an individual appears in a data system is recorded. The result is a frequency distribution for a given period of time, 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 individuals 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 (Blumenthal, Dahiya, and Gross 1978). To estimate the size of the population, an incomplete Poisson distribution is fitted 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 **(see exhibit 4).**

The truncated Poisson models have been used for estimating the size of the criminal population from arrest history records (Greene and Stollmack 1981) and the number of persons engaged in using or selling drugs from drug-related arrest data (Woodward, Brecht, and Bonett 1987). 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.

Using a similar rationale, Research Triangle Institute (1988) applied the model to estimate the size of the treatment-susceptible heroin population. Its model was unusual in that two separate 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

Exhibit 4

Truncated Poisson Estimate in an Open Population

In this procedure, one examines events, such as arrests for drug possession, that occur one or more times to individuals in a population during a specific interval of time. The distribution of individuals having various numbers of events is used to extrapolate the size of the population. In the population whose size is to be estimated, some individuals have no events and are never observed, some have one event, some have two, etc. Let n_j be the number of individuals with j events. The population size N is the sum of the n_j, including the n_0 individuals who were never observed.

To estimate N, one fits a distribution to the frequencies $n_1, n_2, n_3, \ldots$, then uses the shape of this distribution to infer N. Two assumptions about the events suffice to determine their distribution: first, they occur randomly over both individuals and times, and second, their rate is homogeneous across the population. Under these assumptions, the number of counts have a Poisson distribution, with unknown rate parameter λ,

$$P_j = \frac{e^{-\lambda}\lambda^j}{j!}, \quad j = 0, 1, 2, \ldots$$

The mean number of events of each type is

$$\mu_j = NP_j = N\frac{e^{-\lambda}\lambda^j}{j!}, \quad j = 0, 1, 2, \ldots$$

When this theoretical description is fitted to the n_j, estimates of N and λ are obtained. Specific estimation procedures are given by Blumenthal, Dahiya, and Gross (1978). Once the distribution has been fitted, the agreement of the theoretical Poisson frequencies $\hat{\mu}_j$ and the observed data n_j can be tested with the X^2 or G^2 statistics described for the incomplete tables. The key assumptions underlying this procedure are those of the Poisson event model: that the population is homogeneous, that the events are independent, and that their rate is constant in time.

Prevalence Estimation Methods (continued)

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 upon 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 of arrest may increase as the offender becomes known to the police. The effects of such violations of independence 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*, which assess the size and character of a biological population based on repeated marked samples, have been developed and extensively analyzed (for reviews see Seber 1982, 1986). The most common class of ecological sampling models that are 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 (Jolly 1965, 1982; Seber 1965, 1982) and a related model by Cormack (1979, 1981, 1985).

However, a number of the assumptions of these models are unrealistic for drug-using populations, and none of these models has yet been applied to drug-use prevalence estimation. For example, the Jolly-Seber model assumes that (1) all samples are of independent and identical capture probability and survival probability, (2) samples are instantaneous and release is made immediately after each sampling, (3) the drug user's behavior is unaffected by the capture history, and (4)

there is no temporary emigration from the population. These assumptions can hardly be satisfied by existing data or the population characteristics of drug abusers.

Another variety of the open-population multiple-capture approach developed recently is a *Markov-based dynamic recapture model* (Wickens 1990). 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 in which users are initially drawn from an infinite population of nonusers. After this first observation, the balance of the process is governed by the dynamics of a state structure that 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 (Wickens 1982). The full history probabilities are therefore the product of three terms: (1) the size of the sample in which an individual is first observed; (2) the probability of that observation; and (3) the probability of the observation history subsequent to the initial observation, characterized by the full variety of such histories. This procedure generates estimates of the size of the population from which the observations are drawn **(see exhibit 5).**

This Markov estimation technique has the advantage of being able to provide a somewhat more realistic dynamic description of the drug-using process. Its weakness is that the model comes with some strong intrinsic statistical assumptions of homogeneity and independence of history. In addition, to be practically applied, the model can have only a minimal dynamic structure. The limitation on the complexity of the model is necessary if its parameters are to be identifiable and estimated. A rich data source is needed to identify any complex structure, and such comprehensive data are not usually available.

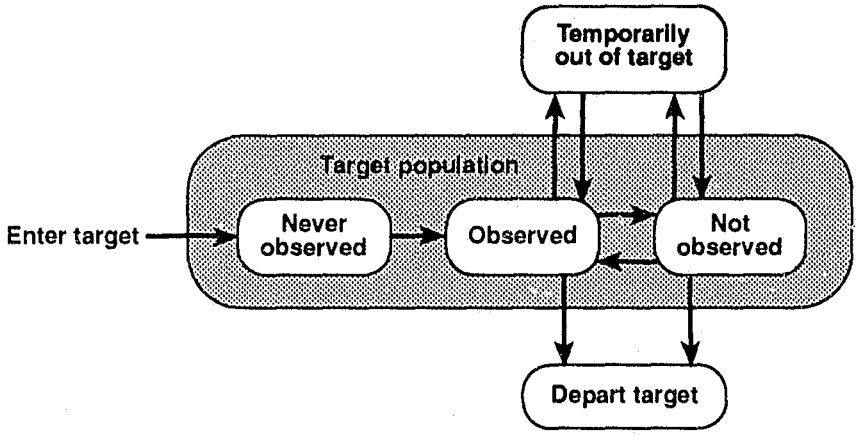
Although the open-population models potentially allow a more realistic picture of drug-using populations as they evolve over time than do closed-population models, such models still require certain restrictive assumptions. Some of these assumptions are particular to the specific models developed in ecological situations, and others serve to simplify the required statistical models so that parameters can be adequately estimated. The latter assumptions include the requirements that individuals behave independently of each other and that the model's parameters are homogeneous over the population. A number of applications have shown that these assumptions are often violated in biology and health sciences (e.g., Huber 1962; Manly 1971; Wittes 1974), and simulation studies have shown potentially large bias in population estimates under such conditions (e.g., Carothers 1973; Gilbert 1973).

Exhibit 5

Markov and Semi-Markov Open-Population Models

In this procedure, the data from a repeatedly observed population are treated as sampling histories of individuals. As in the closed-population multiple-capture models, a probabilistic model is fitted to the observed histories, then extrapolated to count the unobserved members of the population. The Markov and semi-Markov models describe the way that the individuals change between observations as transitions among a small set of discrete states. Individuals enter into the target population, pass through various states of observation and nonobservation, and possibly may leave the target population. In a semi-Markov process, the individual's passage among the states is modeled in continuous time by a homogeneous transition process. For data from a series of samples, the transitions are only observed at the times of sampling, and a simpler discrete-step Markov chain is used.

The models for population-size estimation use a small number of states to describe the individuals. For example, a model of drug use might be formulated from the states that flow from one to another according to the pattern



The circled cells correspond to the states of the individual drug users. The number of individuals in each state must be counted. Note that individuals in the **Never observed** state have never been recorded, while those in the **Not observed** state have been seen at least once before, although not in the current sample.

The essential feature of a Markovian process is the assumption that the current state of an individual provides all available information about the individual's future behavior. How the individual initially reached that state is irrelevant. This independence of history means that the properties of the process are summarized by a set of transition rules pertaining to the individual states. For example, individuals in the **Not observed** state are characterized by the three probabilities (or rates) that indicate how likely an individual in that state is to pass to each of the three states to which that state is connected.

The transition probabilities among the observation and postobservation states can be estimated from the observed histories and are used to infer the target-population size. To estimate the size of the **Never observed** state, one assumes that the rate at which individuals enter the **Observed** state from this state is the same as that at which they enter the **Observed** state from the **Not observed** state. The essence of the argument runs as follows. Suppose that π is the probability of entering the observed state, that N_j is the size of the unobserved population, and that n_j individuals are seen for the first time in the jth sample. On average, n_j should equal $\pi N j$. Using an estimate $\hat{\pi}$ of π that is to be obtained from the observed individuals, the number of unobserved individuals is estimated to be

$$\hat{N} = \frac{n_j}{\hat{\pi}}$$

Similar procedures are used to estimate the size of the out-of-target sets. The size of the target population is the sum of the sizes of its three component states. Some constraints on the transition probabilities are necessary to obtain unambiguous estimates.

The crucial assumption underlying these methods is that the dynamics of unseen individuals match those of the individuals who have been observed. Only in this way can the parameters be correctly transferred back to obtain $\hat{N}j$. Moreover, the transition structure must adequately approximate the dynamics of real individuals. It should be recognized that these models impose a considerable structure on the data and cannot be accurately used without a history sequence of some length.

Theoretically, surveys or other data sources used by multiple-capture methods (either closed- or open-population) 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 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 linking information is difficult to obtain.

Because the multiple-recapture models attempt only to estimate the number of unobserved individuals from the observed sample, they cannot extrapolate to completely unobserved subpopulations. Like any of the statistical methods, they cannot completely solve the hidden-population problem as it was defined earlier. However, open-population models generally are based on probability theory and have better known statistical properties that allow estimates to be evaluated by standard methods such as confidence intervals. In addition, boundaries of generalizability can usually be inferred by data coverage and model specifications.

System Dynamics Modeling

System dynamics is a general methodology, first developed in the late 1950's, for analyzing dynamic phenomena through the use of simulation models based on information-feedback control theory. A system dynamics model consists of an interconnected set of difference equations representing continuous-time movement and accumulations of people, materials, and information. After being assigned initial conditions consistent with historical data, this set of equations is used to generate output over time. If the model is a valid one, this output will closely mimic the true course of events, and the model may be used for prevalence estimation and for making conditional forecasts **(see exhibit 6).**

System dynamics models typically attempt to explain observed dynamics as being the consequence of endogenous feedback relationships among constituent variables. This endogenous perspective distinguishes the system dynamics approach from other modeling methods discussed in this paper, such as synthetic estimation, which depend heavily 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 contain a greater variety of system variables than the multiple-capture or Markov-type models, and tend to be larger in scope. This enlarged scope is rendered manageable by modeling flows as aggregate measures, rather than by keeping a unique record of every individual unit in the flow, as in the multiple-capture models.

Exhibit 6

System Dynamics Models

In this technique, quantities such as the prevalence of drug use are measured by embedding them in a larger model that describes the feedback relationships among these quantities and other social and economic factors. A system dynamics model describes continuous-time change in quantities and factors through an interconnected set of differential equations. System states, such as the number of individuals in a particular state of drug use, are represented by *level variables L(t)*. The level variables are related to one another and change according to *net-rate functions f*; for each level **L** there is a single net-rate function $\mathbf{f}_L$. These net-rate functions utilize intermediate variables known as *rates* and *auxiliaries*. Rate variables define identifiable flows into and out of the levels, such as drug use initiation, escalation, quitting, and relapse. Auxiliary variables define other concepts that make the model more intuitive and natural, such as retail price and perceived health risk. The net-rate functions are parameterized by *constants C*, which include delay times and strength-of-response parameters, such as price elasticities. The net-rate functions may also be influenced by *exogenous variables X(t)* that change in time and are determined by external data, such as the baseline population growth rate and the fraction of drug imports that are seized. With the many level and exogenous variables, constants, and net-rate functions written as vectors and denoted in boldface, a system dynamics model is expressed as

$$\frac{dL(t)}{dt} = f_L(t) = f[L(t), X(t); C]$$

When investigating the model, these differential equations are not solved analytically, but are expressed as difference equations,

$$L_t = L_{t-1} + dtf(L_{t-dt}, X_{t-dt}, C)$$

for small time intervals ***dt***. The variables are simulated starting at some initial time such as $t = 1976$ and stepping forward to subsequent times. The computation interval ***dt*** has no real-world significance and is chosen to be sufficiently small such that the simulation output lies close to the true solution of the corresponding differential equations.

The validation of a system dynamics model, unlike that of the simpler probabilistic models described elsewhere, is an ongoing process of building confidence in the realism of the model's structure and behavior, using a variety of largely qualitative tests (Forrester and Senge 1980). The complexity of the models allows for similar observed behavior to be generated by a variety of structures, which must be carefully chosen if unobserved dynamic quantities, such as prevalence, are to be accurately described. Confidence in the model and its estimates is enhanced when all equations have concrete real-life significance, are dimensionally correct, and operate appropriately even under extreme conditions. Confidence is also enhanced when the model faithfully re-creates the dynamic patterns and correlations observed in real life, and when it brings to light behavior in the real system that has gone unrecognized or unexplained (Mass 1991).

Prevalence Estimation Methods (continued)

Examples of system dynamics modeling studies of illicit drug use include the "Persistent Poppy" model of Levin, Roberts, and Hirsch (1975), models developed by Gardiner and Shreckengost (1985, 1987; Shreckengost 1984, 1985), and a recent prevalence model developed by Homer (1990).

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 treatment programs, it was developed at a time when the numerical data needed for its calibration and validation were lacking.

The models developed by Gardiner and Shreckengost address drug supply and demand on a national level, with specific application to heroin and cocaine. They have been used primarily to make inferences about drug import levels, thus focusing on quantity measures rather than persons or events. The model developed for heroin was adapted to estimate the number of users, but such estimates were shown to be rather sensitive to uncertain assumptions about the number and type of drug-user categories (Shreckengost 1984). Also, the models of Gardiner and Shreckengost may lack sufficient feedback structure and internally generated momentum to be useful for prevalence estimation and forecasting.

A recent national model of cocaine use (Homer 1990) was developed to make inferences about unobserved populations from available data sources related to drug-use prevalence. This model reproduces historical drug-indicator data from 1976 onward and produces prevalence estimates and forecasts for several population categories, including casual and compulsive users. For this purpose, the model utilizes two sorts of information.

First, it uses information about the logical relationships among various population categories. For example, the model depicts the escalation process by which casual users become compulsive users, and distinguishes the effects of powder cocaine from those of crack. Second, the model uses information about the logical relationships between the population categories and other indicators, including morbidity and mortality, drug-related arrests, retail price and purity, etc. For example, the model is calibrated to reflect the idea that compulsive users are more likely than casual users to show up in DAWN data on cocaine-related morbidity and mortality. With sufficient numerical data and knowledge of a phenomenon's structure and dynamics, the range of estimates for hidden populations may be narrowed to a considerable degree.

System dynamics is an attractive approach for prevalence estimation largely because it can be used to explain history, fill in the gaps in indicator data, and project outcomes under different assumed scenarios and policy interventions. In addition, a system dynamics model can often help detect certain flaws in existing indicator data sets, such as possible logical inconsistencies between incidence data and prevalence data.

However, the very flexibility of system dynamics opens the door to potential model misspecification, a danger that becomes greater as the number of variables and conceptual scope of the model increase relative to the quantity of relevant data. Also, like any method, the accuracy of a system dynamics model is sensitive to the quality of the data used to calibrate it. Nevertheless, a well-specified system dynamics model may be useful even if it falls short in numerical precision because it reveals and anticipates trends that other methods may miss.

Because of the complexity of system dynamics models and the risks of misspecification, different tests for building confidence into such models have been offered that go well beyond the usual requirement that historical data be reproduced (Forrester and Senge 1980). 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 pitfalls in its application that one must be careful to avoid.

Prevalence Estimation Methods (continued)

Summary

None of the prevalence estimation methods, whether reviewed here or known to the field, can provide estimates without knowledgeable and careful application or without conditional limitations. Some of these limitations are due to the necessary simplicity of the assumptions of the model, while others are due to the demands for specific data of a certain level of quality.

Static models typically take "snapshots" of the drug-use problem of a specified time or area, and some give additional descriptions in terms of demographic distribution. Dynamic models provide prevalence estimates for time-related processes; they may also specify parameters of the processes and offer some forecasting capability.

Synthetic estimation relies on prevalence estimates from independent sources (calibration samples) to extrapolate from and provide estimates not otherwise available for the desired geographic areas. These independent estimates, in most cases, have come from surveys that are subject to numerous criticisms. Multiple-capture models (in closed or open populations) provide estimates of an incompletely observed population by projecting from the capture pattern of samples observed over time and without requiring independent estimates. However, data of sufficient richness to support this type of application 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 influence the prevalence of drug use and its change over time. However, the building, calibration, and validation of a system dynamics model—especially one that attempts to represent in detail a relatively complex phenomenon—are generally difficult and require special expertise and caution.

The strengths and limits of these various methods are summarized **(see 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 the same as its weakness.

For example, although system dynamics modeling provides a broad framework to examine the phenomena and is able to utilize data from many sources, the data available to specify the model are frequently insufficient to fully resolve the complexity of the structure these models contain. System dynamics modeling is also an excellent tool for making projections and for answering policy-simulation, or "what if," questions.

Open-population multiple-capture models focus on a minimal dynamic structure, a focus which tends to unrealistically simplify the processes involved, often using, for example, only those data concerned with treatment admissions. However, because these multiple-capture models are based on probability theory, their

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. Makes 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 timeframe, 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.	Requires specific form/type of data (e.g., matching individuals across data sources). Dependence on simplified probability model of independent and identical observations.
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.	Requires 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 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.

Prevalence Estimation Methods (continued)

statistical properties are well-defined. When appropriately applied, confidence in the estimation results can be judged by conventional confidence intervals and other statistical tests.

Synthetic estimation uses prevalence figures from the carefully studied calibration sample. It is useful because local planning must, out of necessity, rely on such estimates in the absence of good local data. Because independent local surveys are generally expensive and are not often feasible for local agencies to conduct, synthetic estimation fulfills a clear need. However, the quality of these estimates entirely depends upon the quality of the available indepth data, and a direct mapping to another area or population is not always appropriate.

These methods also vary in the number of assumptions required; the more complicated models make many more, and more carefully qualified, assumptions. Assumptions involved in some of the simpler models, such as population projections, are usually explicit and thus readily subject to examination. By contrast, in a typical application of system dynamics modeling to drug abuse—which usually includes a wide variety of variables—many assumptions are made, and these assumptions are usually complexly interconnected and are often not readily discernible. The adequacy and effects of these assumptions must be subject to careful scrutiny.

Different types of data are best suited for only certain types of estimation methods, 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 also use event-based and quantity-based data.

In selecting a model, one must consider the appropriate use of the data that are available. When data representativeness is not an issue, synthetic estimation methods are the least costly and the 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 timeframe, closed-population models are applicable. Open-population models are used to their best advantage to reflect the dynamic aspect of drug-use progression when individuals are traced over time. System dynamics models are applied when multiple sources of indicator data about the system over time are available and when the primary interest is understanding the dynamic interrelationships among these indicators within the system.

Concluding Comments

Although numerous difficulties are associated with prevalence estimation in whatever content domain it is attempted, estimates are necessary to make decisions for resource allocation, program planning, and other purposes. Recognizing and understanding these difficulties promote the appropriate use of the available data and provide more defensible prevalence estimations. Development in at least three areas is needed to improve estimation of drug prevalence:

Improving the understanding of drug-use phenomena. Timely and thorough information about illicit drug production, distribution, and consumption characteristics allows more appropriate choice and specification of models that better represent drug use and lead to better estimates. Further, 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.

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 would allow statistical techniques to yield high-quality, comprehensive estimates. Existing data systems have problems primarily with missing individuals who are 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 (CODAP data) have had recent emergency room care for drug-related health problems (DAWN data), and vice versa, or the overlap of arrestees testing positive for drug usage (DUF data) with either of these health or treatment indicators. If important common information were available from all indicator series, integrating these indicators would result in nonoverlapping prevalence estimates and in the specification of more precise, consistent, and functional relationships.

Continued development and improvement of prevalence estimation techniques and their appropriate utilization. Current drug-use prevalence estimation is flawed by the deficiencies of the existing data systems and the 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 to some degree on improving existing or developing new statistical techniques to remedy the data deficiencies.

The techniques discussed in this report have all made their contribution to prevalence estimation. However, because of the complex, dynamic nature of drug use and the restrictive sampling or incomplete data in the existing indicators, no single method can adequately produce estimates for all categories of users. A single

Concluding Comments (continued)

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 phenomena 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 are identified.

In addition, alternative models using different approaches and data sources are necessary to validate one another 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.

References

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

Blumenthal, S., R.D. Dahiya, and A.J. Gross (1978). "Estimating the complete sample size from an incomplete Poisson sample." *Journal of the American Statistical Association*, 73(361):182–187.

Brookmeyer, R., and M.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.

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

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

Chandra-Sekar and W.E. 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.

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 Cooperative Publishing House.

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.

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.

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.

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.

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

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.

References (continued)

Ericksen, E.P., and J.B. Kadane (1985). "Estimating the population in a census year: 1980 and beyond." *Journal of the American Statistical Association*, 80(389):98–109.

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

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

Frank, O. (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, DC: National Academy Press.

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

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

Gardiner, L.K., and R.C. Shreckengost (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, Department of Health and Human Services Publication No.(ADM)85–1402. Washington, DC: U.S. Government Printing Office.

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

Gerstein, D., and H. Harwood (eds.) (1990). *Treating Drug Problems: Volume 1*. Washington, DC: National Academy Press.

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

Graham, D. (1988). *Detailed Diagnoses and Procedures for Patients Discharged from Short-Stay Hospitals*. Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Services, National Center for Health Statistics.

Greene, M.A., and S. Stollmack (1981). "Estimating the number of criminals." In J.A. Fox (ed.), *Models in Quantitative Criminology*, pp. 1–24. NY: Academic Press.

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.

Hamill, D. (1988). "Indicator Approach." Paper presented at the Prevalence Estimation Technique Conference, Washington, D.C., November.

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, Department of Health and Human Services Publication No. (ADM) 85–1402. Washington, DC: U.S. Government Printing Office.

Harrison, L.D. (1990). *The Validity of Self-Reported Drug Use Among Arrestees.* Washington, DC: National Institute of Justice.

Homer, J. (1990). "A System Dynamics Simulation Model of Cocaine Prevalence." Final Report to the National Institute of Justice, Grant 87–IJ–CX–0042.

Huber, J.J. (1962). "Trap response of confined cottontail populations." *Journal of Wildlife Management*, 26(2):177–185.

Johnston, L.D., P.M. O'Malley, and J.G. Bachman (1988). *Use of Licit and Illicit Drugs by America's High School Students 1975–1984*, pp. 34–37. Rockville, MD: U.S. Department of Health and Human Services.

Johnston, L.D., P.M. O'Malley, and J.G. Bachman (1989). *Drug Use, Drinking, and Smoking: National Survey Results from High School, College, and Young Adults Populations 1975–1988.* Rockville, MD: U.S. Department of Health and Human Services.

Jolly, G.M. (1965). "Explicit estimates from capture-recapture data with both death and immigration—stochastic model." *Biometrika*, 52:225–247.

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

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, MA: Lexington Books.

Khalsa, H., M.D. Anglin, and A. Paredes (1991). "Natural history of cocaine addiction." In NIDA Research Monograph *Advances in Cocaine Treatment.*

Levin, G., E.B. Roberts, and G.B. Hirsch (1975). *The Persistent Poppy*, pp. 229. Cambridge, MA: Ballinger.

Levy, P.S. (1979). "Small area estimation—synthetic and other procedures, 1968–1978." In J. Steinberg (ed.), *Synthetic Estimates for Small Areas.* NIDA Research Monograph 24. Department of Health, Education, and Welfare, Washington, DC: U.S. Government Printing Office.

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

References (continued)

Mass, N.J. (1991). "Diagnosing surprise model behavior: A tool for evolving behavioral and policy insights (1981)." *System Dynamics Review*, Archives section, 14:209–228.

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

Nelson, C. (1988). *The National Ambulatory Medical Care Survey: United States, 1975–1981 and 1985 Trends*. Hyattsville, MD: U.S. Department of Health and Human Services, Public Health Services, Center for Disease Control, National Center for Health Statistics.

ONDCP (Office of National Drug Control Policy). (1990). *National Drug Control Strategy*, The White House.

O'Malley, P.M., J.G. Bachman, and L.D. Johnston (1984). "Reliability and consistency in self-reports of drug use." *International Journal of the Addictions*, 18:805–824.

Office of Technology Assessment. (1990). *The Effectiveness of Drug Abuse Treatment: Implications for Controlling AIDS/HIV Infection*. Background paper on AIDS-related issues. OTA–BP–H–73. Washington, DC: U.S. Government Printing Office.

Person, P.H., R.L. Retka, and J.A. Woodward (1976). *Toward a Heroin Problem Index—an Analytical Model for Drug Abuse Indicators*. Department of Health, Education, and Welfare Publication Nos. (ADM) 176–367 and (ADM) 78–367. Washington, DC: U.S. Government Printing Office.

Person, P.H., R.L. Retka, and J.A. Woodward (1977). *A Method for Estimating Heroin Use Prevalence*. Department of Health, Education, and Welfare Publication No. 017–024–00589–4. Washington, DC: U.S. Government Printing Office.

Petersen, C.G.J. (1894). "On the biology of our flat-fishes and on the decrease of our flat-fish fisheries with some observations showing how to remedy the latter and promote the flat-fish fisheries in our seas east of the Skaw." *Report of the Danish Biological Station*, IV:1893–4.

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

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.

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

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

Seber, G.A.F. (1982). *The Estimation of Animal Abundance and Related Parameters (Second Edition)*. London: Griffin.

Seber, G.A.F. (1986). "A review of estimating animal abundance." *Biometrics*, 42:267–292.

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

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

Siegel, R.K. (1984). "Changing patterns of cocaine use: Longitudinal observation, consequences, and treatment." In J. Grabowski (ed.), *Cocaine: Pharmacology, Effects, and Treatment of Abuse*, pp. 92–110. National Institute on Drug Abuse Research Monograph 61. Washington, DC: U.S. Government Printing Office.

U.S. Senate, Committee on the Judiciary (1990). *Hard-Core Cocaine Addicts: Measuring—and Fighting—the Epidemic*. Staff report, May 10, 1990.

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

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

Wickens, T.D. (1989). *Multiway Contingency Tables Analysis for the Social Sciences*. Hillsdale, NJ: Lawrence Erlbaum Association.

Wickens, T.D. (1990). "Multiple-Capture Estimation of Population Size With Dynamic State Change." Manuscript in preparation.

Wish, E.D. (1990). "U.S. drug policy in the 1990s: Insights from new data from arrestees." *The International Journal of the Addictions*, 25(3A):377–410.

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

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. NIDA Research Monograph 57, Department of Health and Human Services Publication No. (ADM) 85–1402. Washington, DC: U.S. Government Printing Office.

References (continued)

Woodward, J.A., L. Brecht, and D.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.

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.

Worldwide Cocaine Situation: 1990 (1991). U.S. Department of Justice, Drug Enforcement Administration, Office of Intelligence.

The Assistant Attorney General, Office of Justice Programs, establishes the policies and priorities, and manages and coordinates the activities of the Bureau of Justice Assistance, Bureau of Justice Statistics, National Institute of Justice, Office of Juvenile Justice and Delinquency Prevention, and the Office for Victims of Crime.

U.S. Department of Justice
Office of Justice Programs
National Institute of Justice

Washington, D.C. 20531

Official Business
Penalty for Private Use $300

BULK RATE
POSTAGE & FEES PAID
DOJ/NIJ
Permit No. G–91

ec niques for the | 's ima ion o ici rug-Use Prevalence

National Institute of Justice